

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 14:32:48 ; Search time 6517 Seconds  
(without alignments)  
11136.056 Million cell updates/sec

Title: us-09-854-300-7

Perfect score: 1774

Sequence: 1 agctgagagctccacgcgcgt.....atagcgacgtctctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: gb Da:\*  
2: gb Htg:\*  
3: gb In:\*  
4: gb Om:\*  
5: gb Ov:\*  
6: gb Pat:\*  
7: gb Pl:\*  
8: gb Pl:\*  
9: gb Pr:\*  
10: gb Ro:\*  
11: gb Sts:\*  
12: gb Sy:\*  
13: gb Un:\*  
14: gb Vi:\*  
15: em Ba:\*  
16: em Fun:\*  
17: em Hum:\*  
18: em In:\*  
19: em Mu:\*  
20: em Om:\*  
21: em Or:\*  
22: em Ov:\*  
23: em Pat:\*  
24: em Ph:\*  
25: em Pl:\*  
26: em Ro:\*  
27: em Sts:\*  
28: em Un:\*  
29: em Vi:\*  
30: em Htg Hum:\*  
31: em Htg Inv:\*  
32: em Htg Other:\*  
33: em Htg Mus:\*  
34: em Htg Pln:\*  
35: em Htg Rod:\*  
36: em Htg Mam:\*  
37: em Htg Vtl:\*  
38: em Sy:\*  
39: em Htgo Hum:\*  
40: em Htgo Mus:\*  
41: em Htgo Other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 1662.8 | 93.7        | 2796   | 6 AX337968  | AX337968 Sequence  |
| 2          | 1659.6 | 93.6        | 2830   | 6 AK074264  | AK074264 Homo sapi |
| 3          | 1431.4 | 80.7        | 2876   | 6 AX576252  | AX576252 Sequence  |
| 4          | 1280.8 | 72.2        | 1297   | 9 AF394689  | AF394689 Homo sapi |
| 5          | 1205.2 | 67.9        | 1928   | 10 BC010477 | BC010477 Mus muscu |
| 6          | 1203.6 | 67.8        | 1936   | 10 AB041548 | AB041548 Mus muscu |
| 7          | 1137.2 | 64.1        | 1513   | 10 AY112656 | AY112656 Mus muscu |
| 8          | 1106.8 | 62.4        | 1287   | 10 AF426411 | AF426411 Mus muscu |
| 9          | 714.4  | 40.3        | 1871   | 9 AK027169  | AK027169 Homo sapi |
| 10         | 674.4  | 38.0        | 169812 | 2 AL391315  | AL391315 Homo sapi |
| 11         | 674.4  | 38.0        | 169812 | 2 AL591467  | AL591467 Homo sapi |
| 12         | 422.2  | 24.2        | 252305 | 2 EX470094  | EX470094 Mus muscu |
| 13         | 422.2  | 23.6        | 239385 | 2 AC109686  | AC109686 Rattus no |
| 14         | 418.2  | 23.6        | 239385 | 2 AC109686  | AC109686 Rattus no |
| 15         | 392.2  | 22.1        | 404    | 6 AX071980  | AX071980 Sequence  |
| 16         | 372.6  | 21.0        | 646    | 6 AX341568  | AX341568 Sequence  |
| 17         | 310    | 17.5        | 435    | 6 AX396961  | AX396961 Sequence  |
| 18         | 280.4  | 15.8        | 1020   | 10 AY155441 | AY155441 Mus muscu |
| 19         | 280.4  | 15.8        | 195962 | 2 AC115936  | AC115936 Mus muscu |
| 20         | 278.4  | 15.7        | 244434 | 2 AC111461  | AC111461 Rattus no |
| 21         | 268.8  | 15.2        | 1391   | 9 AB070023  | AB070023 Macaca fa |
| 22         | 268.8  | 15.2        | 1395   | 9 AB070058  | AB070058 Homo sapi |
| 23         | 266.8  | 15.0        | 1429   | 9 AK098524  | AK098524 Homo sapi |
| 24         | 266.8  | 15.0        | 1430   | 9 BC022038  | BC022038 Homo sapi |
| 25         | 266.8  | 15.0        | 1446   | 9 AF447589  | AF447589 Homo sapi |
| 26         | 266.8  | 15.0        | 120984 | 9 AC006463  | AC006463 Homo sapi |
| 27         | 257.8  | 14.5        | 1148   | 10 AY155440 | AY155440 Mus muscu |
| 28         | 256.4  | 14.5        | 1328   | 9 BC023264  | BC023264 Homo sapi |
| 29         | 254.8  | 14.4        | 1355   | 6 AR306585  | AR306585 Sequence  |
| 30         | 254.8  | 14.4        | 1355   | 6 AX061659  | AX061659 Sequence  |
| 31         | 253.2  | 14.3        | 1261   | 9 AK098654  | AK098654 Homo sapi |
| 32         | 220    | 12.4        | 119995 | 10 AL731648 | AL731648 Mouse DNA |
| 33         | 220    | 12.4        | 252305 | 2 EX470094  | EX470094 Mus muscu |
| 34         | 217.4  | 12.3        | 239385 | 2 AC109686  | AC109686 Rattus no |
| 35         | 209    | 11.8        | 1200   | 6 AX083416  | AX083416 Sequence  |
| 36         | 209    | 11.8        | 1527   | 6 AX083426  | AX083426 Sequence  |
| 37         | 209    | 11.8        | 2001   | 9 BC019355  | BC019355 Homo sapi |
| 38         | 209    | 11.8        | 2005   | 6 BD127528  | BD127528 Primer fo |
| 39         | 209    | 11.8        | 2005   | 9 AK075141  | AK075141 Homo sapi |
| 40         | 209    | 11.8        | 2021   | 9 BC045743  | BC045743 Homo sapi |
| 41         | 209    | 11.8        | 2063   | 6 BD063226  | BD063226 Secreted  |
| 42         | 209    | 11.8        | 2452   | 6 BD127402  | BD127402 Primer fo |
| 43         | 209    | 11.8        | 2452   | 9 AK074985  | AK074985 Homo sapi |
| 44         | 207.4  | 11.7        | 2013   | 9 BC032328  | BC032328 Homo sapi |
| 45         | 201.2  | 11.3        | 2328   | 6 AX644275  | AX644275 Sequence  |

## ALIGNMENTS

| RESULT 1   | AX337968  | 2796 bp     | DNA | linear | PAT 09-JAN-2002 |
|------------|---|-------------|-----|--------|-----------------|
| LOCUS      | AX337968  |             |     |        |                 |
| DEFINITION | Sequence 9 from Patent WO0194391.                             |             |     |        |                 |
| ACCESSION  | AX337968  |             |     |        |                 |
| VERSION    | AX337968.1  | GI:18128679 |     |        |                 |
| KEYWORDS   |   |             |     |        |                 |
| SOURCE     | Homo sapiens (human)  |             |     |        |                 |
| ORGANISM   | Homo sapiens  |             |     |        |                 |
| REFERENCE  | 1   |             |     |        |                 |
| AUTHORS    | Yue H., He A., Nguyen D.B., Yao M.G., Bandman O., Burford N., |             |     |        |                 |
| TITLE      | Tang Y.T., Xu Y., Hafalia A., Azimaei Y. and Walla N.K.       |             |     |        |                 |
|            | Intracellular signaling proteins                              |             |     |        |                 |

JOURNAL Patent: WO 0194391-A 9 13-DEC-2001;  
 Incyte Genomics, Inc. (US)  
 Location/Qualifiers  
 source 1..2796  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 2791668CBI"

BASE COUNT 821 a 548 c 620 g 807 t  
 ORIGIN

Query Match 93.7%; Score 1662.8; DB 6; Length 2796;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 74 CCGAGAGCTGATCTGCGGCAACCTGTGTGTGACCTTACGCTCTCTGCTCGAGC 133  
 DB 28 CCGAGAGCTGATCTGCGGCAACCTGTGTGTGACCTTACGCTCTCTGCTCGAGC 87  
 QY 134 TAGCTGAGCTCCCGAGCTCACTCACTCTCCACCTGAGCGGACCTGCTGAG 193  
 DB 88 TAGCTGAGCTCCCGAGCTCACTCACTCTCCACCTGAGCGGACCTGCTGAG 147  
 QY 194 ACCAGGCTCTGCGCAAGCTGAGAGAGCGGTGCGAGGAGCTGAGGAGCTGAGG 253  
 DB 148 ACCAGGCTCTGCGCAAGCTGAGAGAGCGGTGCGAGGAGCTGAGGAGCTGAGG 207  
 QY 254 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 313  
 DB 208 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 267  
 QY 314 TCCAGATTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373  
 DB 268 TCCAGATTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327  
 QY 374 GGGCGGAGAGCTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433  
 DB 328 GGGCGGAGAGCTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387  
 QY 434 GTGAGACCTGAGAGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493  
 DB 388 GTGAGACCTGAGAGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 447  
 QY 494 GAGCGTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553  
 DB 448 GAGCGTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507  
 QY 554 CCGCACAAGATTTCACGCTGCCACGCTTGGGAGAGCAACCTGCAAGCTCTGCTGG 613  
 DB 508 CCGCACAAGATTTCACGCTGCCACGCTTGGGAGAGCAACCTGCAAGCTCTGCTGG 567  
 QY 614 GCCCTGATCCAGAGCGGCGGCGGCTGACCTTGGCAGAGAGATCACTGCTTATGAG 673  
 DB 568 GCCCTGATCCAGAGCGGCGGCGGCTGACCTTGGCAGAGAGATCACTGCTTATGAG 627  
 QY 674 AGATGGGCGCTGAGAGCGCTGATCTTTAACTTCCCGGAGACCGGCAATGAGTCACTCC 733  
 DB 628 AGATGGGCGCTGAGAGCGCTGATCTTTAACTTCCCGGAGACCGGCAATGAGTCACTCC 687  
 QY 734 ATGCTACCGCGGAG 793  
 DB 688 ATGCTACCGCGGAG 747  
 QY 794 AAAATTCTGCAATCTATTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853  
 DB 748 AAAATTCTGCAATCTATTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 807  
 QY 854 AAAAGAGGCGCTTGGTGAATCACTATCAATTTTCTGCTTCTGCTTCTTCTTCTT 913  
 DB 808 AAAAGAGGCGCTTGGTGAATCACTATCAATTTTCTGCTTCTGCTTCTTCTTCTT 867  
 QY 914 ATTACGCGCACTGTGGGCTATTATTATCTTTATCTTCTGAGAGAGTCAAGATGCA 973

DB 868 ATTACGCGCACTGTGGGCTATTATTATCTTTATCTTCTGAGAGGCTACGAAATGCA 927  
 QY 974 AAGGCTCAAG 1033  
 DB 928 AAGGCTCAAG 987  
 QY 1034 CTTCACATGAG 1093  
 DB 988 CTTCACATGAG 1047  
 QY 1094 GCTGTGTCATGATATTATTAACCAATGATTTGGAACGATTTTACCTGCAACAT 1153  
 DB 1048 GCTGTGTCATGATATTATTAACCAATGATTTGGAACGATTTTACCTGCAACAT 1107  
 QY 1154 ATTTTCATTAAG 1213  
 DB 1108 ATTTTCATTAAG 1167  
 QY 1214 AATGTGACATTAACAGCTTGGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273  
 DB 1168 AATGTGACATTAACAGCTTGGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227  
 QY 1274 TTTCAGATCCCTGATATTCATTAATCTAATAGGCTCCCTCCAGAGAGAGATAT 1333  
 DB 1228 TTTCAGATCCCTGATATTCATTAATCTAATAGGCTCCCTCCAGAGAGAGATAT 1287  
 QY 1334 CGAG 1393  
 DB 1288 CGAG 1347  
 QY 1394 GAGGAG 1453  
 DB 1348 GAGGAG 1407  
 QY 1454 GTGGCAG 1513  
 DB 1408 GTGGCAG 1467  
 QY 1514 AATCAAG 1573  
 DB 1468 AATCAAG 1527  
 QY 1574 ACCATTAGATTAACAG 1633  
 DB 1528 ACCATTAGATTAACAG 1587  
 QY 1634 TTTCATTAATTAAG 1693  
 DB 1588 TTTCATTAATTAAG 1647  
 QY 1694 AATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1753  
 DB 1648 A--TGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1704  
 QY 1754 CATAGGCAAGTTCTCTCAG 1774  
 DB 1705 CATAGGCAAGTTCTCTCAG 1725

RESULT 2  
 AK074264 2830 bp mRNA linear PRI 15-FEB-2002  
 LOCUS AK074264  
 DEFINITION Homo sapiens cDNA FLJ33684 f1s, clone HEP09821.  
 ACCESSION AK074264  
 VERSION AK074264.1 GI:18676818  
 KEYWORDS oligo capding; f1s (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE  
 1 Kanabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,  
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Odayashi, M., Nishi, T.,

Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
 NEDO human cDNA sequencing project  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2830)  
 AUTHORS Sugano, S., Suzuki, Y., Oca, T., Obayashi, M., Nishi, T., Isogai, T.,  
 Shibahara, T., Tanaka, T. and Nakamura, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-PEB-2002) Sunio Sugano, Institute of Medical Science,  
 University of Tokyo, Laboratory of Genome Structure, Human Genome  
 Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan  
 (E-mail: chnals@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,  
 Fax: 81-3-5449-5416)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan: cDNA full insert sequencing;  
 Research Association for Biotechnology: cDNA library construction;  
 5'- & 3'-end one pass sequencing; Department of Virology and Human  
 Genome Center, Institute of Medical Science, University of Tokyo  
 (partly supported by Science and Technology Agency).  
 FEATURES  
 source  
 Location/Qualifiers  
 1..2830  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HEP09821"  
 /cell\_line="HepG2"  
 /cell\_type="hepatoma"  
 /clone\_lib="HEP"  
 /note="Cloning vector: pME18SFL3"  
 251..1537  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB85033.1"  
 /db\_xref="GI:18676819"  
 /translation="MGPPGAGVSCRCGGFSRLTAMCFILALSPQAPSGRAEAWNT  
 AYLAWSRVPHTGVNRITWELSEGVYGDSPLEBAVAVLPDGGALANACPTNIF  
 TPLPWSSTVOVSLALIORGGCTEFADKHLAYRGAAGAVINPGRNINPMISH  
 EGAVIDYALIMLNGKGTILIOSIORGIQVMTVEYKGGWVNHYSIFPVSFPII  
 TAAVYGFYFARLRNARNAOSRKRQKADAKKALIRLDELTKOGKKEIFSPGDS  
 CAVCIPLKYNLDIRILINCHIPIKTCVDPMLHRTCKPMCKDILKALGIVDVEDG  
 SVSLQVRETSNLSAASHEDNRSFTASGTAISVQGTDPPELPHVQSTNLSLQVH  
 HEANSVAVDVIPIVNDPFEEDTPEVNETVAREIKS"  
 BASE COUNT 827 a 564 c 630 g 809 t  
 ORIGIN  
 Query Match 93.6%; Score 1659.6; DB 9; Length 2830;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1688; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

422 GTGAACCTGACGTGTGGAGACTGAGGAGAGAGGCGTGTACGGCCAGACTGCGCTG 481  
 494 GAGCCTGTGGCTGGAGCTTCGTGTACCGCCGACG33CCCG33GCTTAAACGCTGTAA 553  
 482 GAGCCTGTGGCTGGAGCTTCGTGTACCGCCGACG33CCCG33GCTTAAACGCTGTAA 541  
 554 CGGACACGAAATTTCACGCTGCCACGCTTGGAGAGACACGTCGAAGCTCTGGTTG 613  
 542 CGGACACGAAATTTCACGCTGCCACGCTTGGAGAGACACGTCGAAGCTCTGGTTG 601  
 614 GCCCTATCCACACGCGCGCGCGCTGACCTTGCAGACAAAGATCCATCTGGCTTAA 673  
 602 GCCCTATCCACACGCGCGCGCGCTGACCTTGCAGACAAAGATCCATCTGGCTTAA 661  
 674 AGATGGGCGTGTGGAGCGGTACCTTAACTTCCCGGGACCGGACATAGGATCATCCC 733  
 662 AGATGGGCGTGTGGAGCGGTACCTTAACTTCCCGGGACCGGACATAGGATCATCCC 721  
 734 ATGTCTACCCGGGTGCAATGACATTTGTCGATCATGATCGCATCTGAAGGACACA 793  
 722 ATGTCTACCCGGGTGCAATGACATTTGTCGATCATGATCGCATCTGAAGGACACA 781  
 794 AAATTCGACATCTATTCAAAGAGGACATCAAGTGAACAATGTCATTAAGAGGAAA 853  
 782 AAATTCGACATCTATTCAAAGAGGACATCAAGTGAACAATGTCATTAAGAGGAAA 841  
 854 AAACATGGCCCTTGGGTGAATCACTAATCAATTTTGGTTGTTGCTCTTTTAT 913  
 842 AAACATGGCCCTTGGGTGAATCACTAATCAATTTTGGTTGTTGCTCTTTTAT 901  
 914 ATTACGGGCAACCTGGGCTAATTTATCTTTATCTGCTGCAAGGCTACGGAATGA 973  
 902 ATTACGGGCAACCTGGGCTAATTTATCTTTATCTGCTGCAAGGCTACGGAATGA 961  
 974 AGAGCTCAAGCAGAGCAGAGGCAATTAAAGCAGATCTAAAGAACTATTGAAAG 1033  
 962 AGAGCTCAAGCAGAGCAGAGGCAATTAAAGCAGATCTAAAGAACTATTGAAAG 1021  
 1034 CTTCACATGACGACACTGAAACAGAGGACAAAGAAATGGCCCTGATGAGATATTGT 1093  
 1022 CTTCACATGACGACACTGAAACAGAGGACAAAGAAATGGCCCTGATGAGATATTGT 1081  
 1094 GCTGTGTCAATTAATTTGATTAACCAATGATTGTGACGATCTTAACGTGCAACAT 1153  
 1082 GCTGTGTCAATTAATTTGATTAACCAATGATTGTGACGATCTTAACGTGCAACAT 1141  
 1154 ATTTTCATTAAGCAGATGTGTGAACCATGGCTGTAAACCAAGACTGCCCATGTGC 1213  
 1142 ATTTTCATTAAGCAGATGTGTGAACCATGGCTGTAAACCAAGACTGCCCATGTGC 1201  
 1214 AAATGTGACATTAATCTTAAGCTTTGGGAATTAAGGTGATTTGAAGTGAATGATGTCT 1273  
 1202 AAATGTGACATTAATCTTAAGCTTTGGGAATTAAGGTGATTTGAAGTGAATGATGTCT 1261  
 1274 TTAACAAGCCCTGTATTCACAAATTAATCTTAAGTGTCTCTCTCCATGAGAGTAAT 1333  
 1262 TTAACAAGCCCTGTATTCACAAATTAATCTTAAGTGTCTCTCTCCATGAGAGTAAT 1321  
 1334 CCAGACGAGACCGGATCATCTGATATGCTTCACTGACAGAGAACGATGAACGCTCTTG 1393  
 1322 CCAGACGAGACCGGATCATCTGATATGCTTCACTGACAGAGAACGATGAACGCTCTTG 1381  
 1394 GAGGAACACGTGACATCAACAATGAAGTCCTACCTGTGAACCAATGAAGCAATTCCT 1453  
 1382 GAGGAACACGTGACATCAACAATGAAGTCCTACCTGTGAACCAATGAAGCAATTCCT 1441  
 1454 GTGGACATGATGATTATCTCATGATGACCAACCTTTGAAGAGAGCAAACTCTCT 1513  
 1442 GTGGACATGATGATTATCTCATGATGACCAACCTTTGAAGAGAGCAAACTCTCT 1501  
 1514 AATCAAGAGATGCTGTGGAAGATGAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1573

Db 1502 AATCAAGAGACTGCTGTTCGAGAAATTATATCTTAAATCTGTATTAATAGAAACTTGA 1561  
 QY 1574 ACCATTAGTAATTAACAGACTGCAATCAGGGCCTAGTTCTATTAATTAATGGATPA 1633  
 Db 1562 ACCATTAGTAATTAACAGACTGCAATCAGGGCCTAGTTCTATTAATTAATGGATPA 1621  
 QY 1634 TTTAATAAATAAGAGTATGATAGAAAGTGCAGATGACATATATTAAGTAATAGTTAA 1693  
 Db 1622 TTTAATAAATAAGAGTATGATAGAAAGTGCAGATGACATATATTAAGTAATAGTTAA 1681  
 QY 1694 AATGCTTAAAAATATTTAACTGTTAACCTTTTTCACCAAACTCATTAATTAATTTT 1753  
 Db 1682 A--TGGCTTAAAAATTTAACTGTTAACCTTTTTCACCAAACTCATTAATTAATTTT 1738  
 QY 1754 CATAGCAAGTTCTCTCCAG 1774  
 Db 1739 CATAGCAAGTTCTCTCCAG 1759  
 RESULT 3  
 AX576252  
 LOCUS Sequence 33 from Patent WO0162922. 2876 bp DNA linear PAT 08-JAN-2003  
 DEFINITION AX576252  
 AX576252 GI:27646003  
 VERSION AX576252.1 GI:27646003  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Banville,S.C., Greenawalt,L.B., Lincoln,S.E., Stockbreher,T.K.,  
 Amshay,S.C., Chang,S.C., Chen,P.W., Dam,T.C., Liu,T.F., Rosen,B.H.,  
 Russo,F.D., D'Sa,S.A., Spiro,P.A., Bradley,D.L., Chen,A.,  
 Cohen,H.J., Daffo,A., Daniels,S.E., Dufour,G.E., Flores,V.,  
 Fong,W.T., Hodgson,D.M., Jackson,S., Jones,A.L., Panzer,S.R.,  
 Roseberry,A.M., Shah,P., Wright,R.J., Yap,P.B., Yu,J.Y.,  
 Bracher,S.R., Chalup,M.S., Dahl,C.R. and Hillman,J.L.,  
 Polypeptides and corresponding molecules for disease detection and  
 treatment  
 Patent: WO 0162922-A 33 30-AUG-2001;  
 JOURNAL Incyte Genomics, Inc. (US)  
 FEATURES  
 Location/Qualifiers  
 source 1..2876  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: LI:757439.1:2000MAV01"  
 BASE COUNT 840 a 566 c 642 g 825 t 3 others  
 ORIGIN  
 Query Match 80.7%; Score 1431.4; DB 6; Length 2876;  
 Best local Similarity 96.5%; Pred. No. 0;  
 Matches 1670; Conservative 0; Mismatches 28; Indels 33; Gaps 19;  
 QY 74 CCAGAGAGCTGCATCTGCGGCAACCTGTGTGCTGACAGCTACGTCCTCTGAC-TCCGAC 132  
 Db 33 CCAGAGAGCTGCATCTGCGGCAACCTGTGTGCTGACAGCTACGTCCTCTGACCTCCAG 92  
 QY 133 GTAGCTGCAAGCTCCCACTGCTCACTCCATCTCTCCCACTGCGGCGGCGCACTCTCTAA 132  
 Db 93 GTAGCTGCAAGCTCCCACTGCTCACTCCATCTCTCCCACTGCGGCGGCGCACTCTCTAA 152  
 QY 193 GACCAAGGTCCTGCAAGCTGCAAGGAGGCGGTGCCAGGAGGCGCTAGGAACTCGGAG 252  
 Db 153 GACCAAGGTCCTGCAAGCTGCAAGGAGGCGGTGCCAGGAGGCGCTAGGAACTCGGAG 212  
 QY 253 CGCGCGCCCATGAGGCGCGCGCTCGGCGCGGCGCTCTCTGCGCGGCGGCGCTGCGGCTT 312  
 Db 213 CGCGCGCCCATGAGGCGCGCGCTCGGCGCGGCGCTCTCTGCGCGGCGGCGCTGCGGCTT 272  
 QY 313 TTCAGATTGCTGAGTGTGCTCTGCTGCGCCCTGAGTCCGAGGCAACCGGTTCCCG 372

Db 273 TTCAGATTGCTGAGTGTGCTCTGCTGCGCCCTGAGTCCGAGGCAACCGGTTCCCG 332  
 QY 373 GGGGCGCTGAAGAGTGTGAACCGCGTAACTCAAGTGTCTCGGCGGCTTCGCAACCGG 432  
 Db 333 GGGGCGCTGAAGAGTGTGAACCGCGTAACTCAAGTGTCTCGGCGGCTTCGCAACCGG 392  
 QY 433 AGTGAACGTAAGGTGTGAGAGTGAAGAGGAGGCGGTGAAGGCGCAGAGCTCGCGGCT 492  
 Db 393 AGTGAACGTAAGGTGTGAGAGTGAAGAGGAGGCGGTGAAGGCGCAGAGCTCGCGGCT 452  
 QY 493 GAGGCTGTGAGTGTGCTGTGTACCGCTGACCGGCGCGGCGGCTTAAAGCTGTAA 552  
 Db 453 GAGGCTGTGAGTGTGCTGTGTACCGCTGACCGGCGCGGCGGCTTAAAGCTGTAA 512  
 QY 553 CCGGCAACGAAATTCAGAGTGTGCGCAAGGTTGGGGAAGCAACGTCAGTCTGTGGT 612  
 Db 513 CCGGCAACGAAATTCAGAGTGTGCGCAAGGTTGGGGAAGCAACGTCAGTCTGTGGT 572  
 QY 613 GGGCTCATCAACCGCGCGGCGGCGTGCACCTTGCAGACAAAGTCAATCTGGCTTATGA 672  
 Db 573 GGGCTCATCAACCGCGCGGCGGCGTGCACCTTGCAGACAAAGTCAATCTGGCTTATGA 632  
 QY 673 GAAATGGGCGCTGGAAGCGCTCAATCTTAACTCCCGGGAACCGGAATGAAGTCAATCC 732  
 Db 633 GAAATGGGCGCTGGAAGCGCTCAATCTTAACTCCCGGGAACCGGAATGAAGTCAATCC 692  
 QY 733 CATGCTCAACCGGCGTGAAGACATTTGTCATCAATCAATGA-TCCGCAATCTGAAGCA 791  
 Db 693 CATGCTCAACCGGCGTGAAGACATTTGTCATCAATCAATGAAGTGAAGCA 752  
 QY 792 CAAAAATCTGCAATCTTCAAGAGGCAATCAAGTGAAGAGTCAATGAAGTGAAGTGAAG 851  
 Db 753 CAAAAATCTGCAATCTTCAAGAGGCAATCAAGTGAAGAGTCAATGAAGTGAAGTGAAG 812  
 QY 852 AAAAAATGAGGCGCTTGGTGAATCACTATCTATTTTTCG-TTCTGTGCTCTTTT 910  
 Db 813 AAAAAATGAGGCGCTTGGTGAATCACTATCTATTTTTCGTTCTGTGCTCTTTT 872  
 QY 911 ATTATTAAGGCG-GGAATCTGAGGCTATTTTATCTTTATCTGCTGGAAGGCTGAGAA 969  
 Db 873 ATTATTAAGGCGGGAATCTGAGGCTATTTTATCTTTATCTGCTGGAAGGCTGAGAA 932  
 QY 970 TGCAGAGGCTCAAGACAGAGAGAGAGG-CAATTAAGGAGATCTTAAAAAGCTATG 1028  
 Db 933 TGCAGAGGCTCAAGACAGAGAGAGAGGCAATTAAGGAGATCTTAAAAAGCTATG 992  
 QY 1029 GAGGCG-TTCACTACGCACTGGAACAGAGAGCA-GGAATTTGCGCTGAT-GGAG 1085  
 Db 993 GAGGCGTTTCACTACGCACTGGAACAGAGAGCAAGGGAATTTGCGCTGATGGAG 1052  
 QY 1086 ATAGTTGCTGTGCTGATTAATTTGATTAACCAATGATTTGATACGATCTTAAAGT 1145  
 Db 1053 ATAGTTGCTGTGCTGATTAATTTGATTAACCAATGATTTGATACGATCTTAAAGT 1112  
 QY 1146 GCAACCAATTTTCCATTAAGACATGTGTGAACCCATGCTGTTAACACACACACTTGC 1205  
 Db 1113 GCAACCAATTTTCCATTAAGACATGTGTGAACCCATGCTGTTAACACACACTTGC 1172  
 QY 1206 CCATGTCAGAAATGTGACATCTCAAGCTTTGGGAATGAGTGAAGTGAAGTGAAT 1265  
 Db 1173 CCATGTCAGAAATGTGACATCTCAAGCTTTGGGAATGAGTGAAGTGAAGTGAAT 1232  
 QY 1266 CAGTGT-CTTTCAAGTCCCTGT---ATCCAAATGAATCTTAATAGTCTCTCCCA 1320  
 Db 1233 CAGTGTCTTTCAAGTCTTCCCTGATATCAATTAAGAAATATCTAATAGTCTCTCCCA 1292  
 QY 1321 TGAAGGAAATATTCGACAGAGACCGGATCATCTGATATGCTTCACTACAGGAAACGA 1380  
 Db 1293 TGAAGGAAATATTCGACAGAGACCGGATCATCTGATATGCTTCACTACAGGAAACGA 1352  
 QY 1381 TGAACGCGC-TCGAGAGAAACACGTGAGTCAAC-AAATGAATCTTACGCTGAT-AAA 1437  
 Db 1353 TGAACGCGCTTCGAGAGAAACACGTGAGTCAACAAATGAATCTTACGCTGATAAAA 1412

| Query Match | Best Local Similarity | Matches 1885, Conservative                                     | 72.2%<br>99.1%<br>0 | Score 1280.8;<br>Pred. No. 1e-285;<br>Mismatches 12; | DB 9;<br>Indels 0; | Length 1297;<br>Gaps 0; |
|-------------|-----------------------|--|---------------------|--|--------------------|-------------------------|
| QY          | 253                   | CGGCGGCGCGAAGGAGCGCGCGCTGGGCGCGGAGTCTCTGCGCGGAGTGGCGGCTT       | 312                 |  |                    |                         |
| DB          | 1                     | CGGCGGCGCGAAGGAGCGCGCGCTGGGCGCGGAGTCTCTGCGCGGAGTGGCGGCTT       | 60                  |  |                    |                         |
| QY          | 313                   | TTCCAGATGTCGCGCATGAGTGTCTCTGCTGCGCTGAGTCCGAGGACCCGGTCCCG       | 372                 |  |                    |                         |
| DB          | 61                    | TTCCAGATGTCGCGCATGAGTGTCTCTCTGCGCTGAGTCCGAGGACCCGGTCCCG        | 120                 |  |                    |                         |
| QY          | 373                   | GGGGGCTGAAAGCAGTGTGGAACCGCGTACCGGACGTCGCGGAGGTTCCGACACGAG      | 432                 |  |                    |                         |
| DB          | 121                   | GGGGGCTGAAAGCAGTGTGGAACCGGACGTCGCGGAGGTTCCGACACGAG             | 180                 |  |                    |                         |
| QY          | 433                   | AGTGAACCGTACCGGTGTGGAGCTGACGAGAGGAGCGGTGACGAGCACTGCGCGCT       | 492                 |  |                    |                         |
| DB          | 181                   | AGTGAACCGTACCGGTGTGGAGCTGACGAGAGGAGCGGTGACGAGCACTGCGCGCT       | 240                 |  |                    |                         |
| QY          | 433                   | GGAGCGCTGAGGCGCGGAGGTCGTGTACCGGCGCAGCGGCGCGGAGGTCGTAAAGCGCTGTA | 552                 |  |                    |                         |
| DB          | 241                   | GGAGCGCTGAGGCGCGGAGGTCGTGTACCGGCGCAGCGGCGCGGAGGTCGTAAAGCGCTGTA | 300                 |  |                    |                         |
| QY          | 553                   | CCGCGCACAGAAATTCACGAGTCCGACGAGTTTGAGGAGACACCGGTGCAAGTCTCTTGTT  | 612                 |  |                    |                         |
| DB          | 301                   | CCGCGCACAGAAATTCACGAGTCCGACGAGTTTGAGGAGACACCGGTGCAAGTCTCTTGTT  | 360                 |  |                    |                         |
| QY          | 613                   | GGCGCTCATTCACACGCGCGGAGGAGCTGCACTTCGACAGCAAGATTCATCTGGCTATGTA  | 672                 |  |                    |                         |
| DB          | 361                   | GGCGCTCATTCACACGCGCGGAGGAGCTGCACTTCGACAGCAAGATTCATCTGGCTATGTA  | 420                 |  |                    |                         |
| QY          | 673                   | GAGATGGGCGCTGAGAGCGGTCATCTTTAACTTCCCGCGGACCCGCAATGAGTCAATCCC   | 732                 |  |                    |                         |
| DB          | 421                   | GAGAGGGGCGCTGAGAGCGGTCATCTTTAACTTCCCGCGGACCCGCAATGAGTCAATCCC   | 480                 |  |                    |                         |
| QY          | 733                   | CATGTCTCACCGGGGTGAGTGAACATGTTGCAATCATGATGCGCAATCTGAAAGCGAC     | 792                 |  |                    |                         |
| DB          | 481                   | CATGTCTCACCGGGGTGAGTGAACATGTTGCAATCATGATGCGCAATCTGAAAGCGAC     | 540                 |  |                    |                         |
| QY          | 793                   | AAAAAATCTGCAATCTAATCAAGAGGACATCAAGGACATGAGTCACTAGAGTATAGGAA    | 852                 |  |                    |                         |
| DB          | 541                   | AAAAAATCTGCAATCTAATCAAGAGGACATCAAGGACATGAGTCACTAGAGTATAGGAA    | 600                 |  |                    |                         |
| QY          | 853                   | AAAAACATGGCCCTGGGTGAATCACTAATCAATTTTTTTCGTTCTGTGTCCTTTTTAT     | 912                 |  |                    |                         |
| DB          | 601                   | AAAAACATGGCCCTGGGTGAATCACTAATCAATTTTTTTCGTTCTGTGTCCTTTTTAT     | 660                 |  |                    |                         |
| QY          | 913                   | TATTAAGGCGGAGCATGAGGCGTATTTAATCTTTAATCTGTGTGAGAGGCAAGGAATGC    | 972                 |  |                    |                         |
| DB          | 661                   | TATTAAGGCGGAGCATGAGGCGTATTTAATCTTTAATCTGTGTGAGAGGCAAGGAATGC    | 720                 |  |                    |                         |
| QY          | 973                   | AAGAGCTCAAGCAGGAGCAGAGGCAATTAAGGCGAGTGTCTAAAAAGCTATTTGAAG      | 1032                |  |                    |                         |
| DB          | 721                   | AAGAGCTCAAGCAGGAGCAGAGGCAATTAAGGCGAGTGTCTAAAAAGCTATTTGAAG      | 780                 |  |                    |                         |
| QY          | 1033                  | GCTTCAACTTCGCACTGGAACAAGAGACAGAGAAATGGCCCTGATGGAATGTTG         | 1092                |  |                    |                         |
| DB          | 781                   | GCTTCAACTTCGCACTGGAACAAGAGACAGAGAAATGGCCCTGATGGAATGTTG         | 840                 |  |                    |                         |
| QY          | 1093                  | TGCTGTGAGATGGAATGTATAACCAAAATGATTTGGTAAGCATCTTAAAGTGACACA      | 1152                |  |                    |                         |
| DB          | 841                   | TGCTGTGAGATGGAATGTATAACCAAAATGATTTGGTAAGCATCTTAAAGTGACACA      | 900                 |  |                    |                         |
| QY          | 1153                  | TATTTTCATTAAGCATGTGTGACCGATGGCTGTAAAAACAGAACTTGCCCAATGTG       | 1212                |  |                    |                         |

```

Db      901 TATTTCCATAGACAGTGTGTACCCAGCTGTGTAGAACACAGCACTTGCACCAGTG 960
QY      1213 CAATATGACATCTCTCAAGAGCTTTGGCAATTTAGGTGATGTTGAAGATGATCGTCT 1272
Db      961 CAATATGACATCTCTCAAGAGCTTTGGCAATTTAGGTGATGTTGAAGATGATCGTCT 1020
QY      1273 TTACAAAGTCCCTGTATTCATGAAATATCTAATGTAGTCCCTCCCTCCCATGAGAGATTA 1332
Db      1021 TTTACAGTCCCTGTATTCATGAAATATCTAATGTAGTCCCTCCCTCCCATGAGAGATTA 1080
QY      1333 TCGCAGCGAGACCGCATCATCTGATATCTTCACTACAGGGAACAGATGAACCGCTCT 1392
Db      1081 TCGCAGCGAGACCGCATCATCTGATATCTTCACTACAGGGAACAGATGAACCGCTCT 1140
QY      1393 GAGGAGACCGCTGACGTCACAAATGAAGTCTACAGTGTGTAACCAAGAGCAATTC 1452
Db      1141 GAGGAGACCGCTGACGTCACAAATGAAGTCTACAGTGTGTAACCAAGAGCAATTC 1200
QY      1453 TGTGCAAGTGTATTTATCTCATGTGACATGACCAACCACTTTGAGAGAGCAAACTCC 1512
Db      1201 TGTGCAAGTGTATTTATCTCATGTGACCAACCACTTTGAGAGAGCAAACTCC 1260
QY      1513 TAATCAAGAGACTGCTGTGAGAAATTAATCTTA 1549
Db      1261 TAATCAAGAGACTGCTGTGAGAAATTAATCTTA 1297

RESULT 5
LOCUS   BC010477 1928 bp mRNA linear ROD 10-JUN-2003
DEFINITION Mus musculus ring finger protein 128, mRNA (CDNA clone MGC:19395
ACCESSION BC010477
VERSION   BC010477.1 GI:14714672
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
          Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
          Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bat,N.K.,
          Hopkins,R.F., Udalen,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
          Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.W., Hong,L.,
          Scieczka,T.E., Brownstein,M.U., Usdin,T.B., Toshiyuki,S.,
          Carmanici,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
          Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
          McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
          Wooley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
          Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
          Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
          Sanchez,A., Whiting,M., Madan,A., Young,A.C., Rodriguez,Y.,
          Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
          Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
          Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
          Scherch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL   MEDLINE
PUBMED    2238257
1247932
TITLE     2 (bases 1 to 1928)
AUTHORS    Strausberg,R.
JOURNAL    Direct Submission
Submitted (10-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK     MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk

```

```

FEATURES
source
1..1928
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="MGC:19395 IMAGE:3154433"
/tissue_type="mammary tumor, Brca1-/-, MMTV-Cre model, 10
months old, gross tissue."
/clone_lib="NCI CGAP Mam3"
/lab_host="DH10B"
/note="vector: pCMV-Sport6"
1..1928
/gene="Rnf128"
/note="synonyms: 1300002C13Rik, GRAIL"
/db_xref="locusid:66889"
/db_xref="MGI:1914139"
176..1462
/codon_start=1
/product="Riken CDNA 1300002C13"
/protein_id="P4H10477.1"
/db_xref="GI:14714673"
/db_xref="locusid:66889"
/translation="MGPPIGVYGGCGGCAATLAWCFLLATSPHAPSGAEAVWT
AYLVNVRVETGVNRTVSEGVYGDSEVSGVVPDVGALNACNPNTF
TPTVWGSVVOVMALILORGCGTADKILHLSRSGAGVFNPGNRYVPSH
PKAGDYATVMGNKGTILQSTORGLDVTGKKGKGMVNVHVSIFPVASPTI
TAATVGFPIFYSARLNRARQSRKROKAPAKALRKDLRTIKQSPKEIGPDS
CAVCLLVNMDLVNLIKNHIFKTCVPMILHRTCPMKCDILKAGIEVDVDS
SVSLQVFNENKSNVSPHEHSDRSSTASGYASVQADDEPLIEHAGSANTMDLVN
HEANSVAVDVPHVDPNPFEDDETPDQBAVREIKS"
BASE COUNT 482 a 466 c 498 g 482 t
ORIGIN
Query Match 67.9%; Score 1205.2; DB 10; Length 1928;
Best Local Similarity 89.6%; Pred. No. 3,7e-268;
Matches 1364; Conservative 0; Mismatches 149; Indels 10; Gaps 6;
QY 252 GGGCGCGGCATGAGGGGCGCGCGCTGGGGCGGGGCTCTCGCCGCGGTGCGCGCT 311
Db 165 GAGTGGCGGCGCAHGGGGCGCGCGCGGAGTCTACGCGCGGGCTGGGGG 224
QY 312 TTTCAGATTGCTGGGATGCTCTTCTGCTGGGCGCTGAGTCCGAGGACCGGTTCC 371
Db 225 CTGCCCGGCTACGTGGTGTGCTTCTTCTGCTGTGAGTCCGACACCGCGGTTCC 284
QY 372 GGGGGGCTGAAGAGAGTGAACCGGCTACCTCAACGTCCTGGGCGGTTCCGACACG 431
Db 285 GGGGAGCCGAGACCGGTGTGACCGGCTACCTCAACGTCCTGGGCGGTTCCGACACG 344
QY 432 GAGTGAACCGGATGCTGGAGAGCTGAGCAGAGAGGCGCTGTACGACGAGACTCGCGC 491
Db 345 GAGTGAACCGGATGCTGGAGAGCTGAGCAGAGAGGCGCTGTACGACGAGACTCGCGC 404
QY 492 TGGAGCCTGTGGGCTGTGTGACCGCCGACCGGCGCGGCGGCTTAAACCTGTGTA 551

```

Email: cgabs-remail.nih.gov

Tissue Procurement: Lohar Hemphausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HSQC

Web site: <http://www.hpsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gumartne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseg, H.,

Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambavi,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRK Plate: 23 Row: b Column: 12

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 12563610.

Location/Qualifiers

1..1928

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="129, C57BL/6J, FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:19395 IMAGE:3154433"

/tissue\_type="mammary tumor, Brca1-/-, MMTV-Cre model, 10

months old, gross tissue."

/clone\_lib="NCI CGAP Mam3"

/lab\_host="DH10B"

/note="vector: pCMV-Sport6"

1..1928

/gene="Rnf128"

/note="synonyms: 1300002C13Rik, GRAIL"

/db\_xref="locusid:66889"

/db\_xref="MGI:1914139"

176..1462

/codon\_start=1

/product="Riken CDNA 1300002C13"

/protein\_id="P4H10477.1"

/db\_xref="GI:14714673"

/db\_xref="locusid:66889"

/translation="MGPPIGVYGGCGGCAATLAWCFLLATSPHAPSGAEAVWT

AYLVNVRVETGVNRTVSEGVYGDSEVSGVVPDVGALNACNPNTF

TPTVWGSVVOVMALILORGCGTADKILHLSRSGAGVFNPGNRYVPSH

PKAGDYATVMGNKGTILQSTORGLDVTGKKGKGMVNVHVSIFPVASPTI

TAATVGFPIFYSARLNRARQSRKROKAPAKALRKDLRTIKQSPKEIGPDS

CAVCLLVNMDLVNLIKNHIFKTCVPMILHRTCPMKCDILKAGIEVDVDS

SVSLQVFNENKSNVSPHEHSDRSSTASGYASVQADDEPLIEHAGSANTMDLVN

HEANSVAVDVPHVDPNPFEDDETPDQBAVREIKS"





|    | Query Match | Similarity  | 89.5%        | Score | 1203.6     | DB  | 10     | Length | 1396 |
|----|-------------|---|--------------|-------|------------|-----|--------|--------|------|
|    | Best Local  | Similarity  | 89.5%        | Pred. | 0.8        | 6   | 268    |        |      |
|    | Matches     | 1363  | Conservative | 0     | Mismatches | 150 | Indels | 10     | Gaps |
| QY | 252         | GGCGCGCGCCATCGGAGCCGCGCGCTCGGAGCGGAGCTCTCTCGCGCGGTGCTCGGCT      | 311          |       |            |     |        |        |      |
| Db | 176         | GAGTGCACGCGCCAGCGGAGCGCGCGCGCGCGGATCGGAGGTCTACTGCGCGCGCGCTCGGAG | 235          |       |            |     |        |        |      |
| QY | 312         | TTTTCAGATATGCTGGATATGTTGCTTTCGATGGCCCTGATGTCGCGAGGACCGCGGTCC    | 371          |       |            |     |        |        |      |
| Db | 236         | CTGCGCGCTACTGAGCTTGTGATCTTCTCTTGAGCTTGAATTCGACGCGCGCGGTCC       | 295          |       |            |     |        |        |      |
| QY | 372         | GGGGGCGTGAAGAGTGTGGAACCGCGTACTCAACGTGTCCTTGCGGAGTTCCGCAACGG     | 431          |       |            |     |        |        |      |
| Db | 296         | GGGAGACCGAAGCCGTGTGACCGCGTAACTCAACGTGCTCGCGCGGTTCCGACAACG       | 355          |       |            |     |        |        |      |
| QY | 432         | GAGTGAACCGTACGATGTGTGGAGCTGAGAGAGAGAGAGGTGTAAGCGACAGGACTGCGCG   | 491          |       |            |     |        |        |      |
| Db | 356         | GAGTGAACCGACGATGTGTGGAGCTGAGAGAGAGAGAGGTGTAAGCGACAGGACTGCGCG    | 415          |       |            |     |        |        |      |
| QY | 492         | TGGAGCCGTGTGAGTGGAGGCTCTGTATCGCGCCGACGGAGCCCGGAGCGCTTAAACGCTGTA | 551          |       |            |     |        |        |      |
| Db | 416         | TGGAGCCGTGTCTCGGAGGCTCTGTATCGCGCCGACGGAGCCCGGAGCGCTTAAACGCTGTA  | 475          |       |            |     |        |        |      |
| QY | 552         | AACCGCACAGAAATTTCAACGTGACCAAGGTTTGGGAAAGACCGGTGCAAGTCTGTGAT     | 611          |       |            |     |        |        |      |
| Db | 476         | AACCGCACACAAATTTCAACGTGACCAAGGTTTGGGAAAGACCGGTGCAAGTCTGTGAT     | 535          |       |            |     |        |        |      |
| QY | 612         | TGGCCCTCATTCACACGCGCGGAGGCTGCACTTTCGACACAGATTCATCTGCTATG        | 671          |       |            |     |        |        |      |
| Db | 536         | TGGCCCTCATTCACACCGGTGAGGCTGCACTTTCGCGGACAGGTTCACTGCTGCTGAG      | 595          |       |            |     |        |        |      |
| QY | 672         | AAGATGTGGCGGTCTGAGACCGCTCATCTTTAACTTCCCGGAGCCCGCAATGAGATCACC    | 731          |       |            |     |        |        |      |
| Db | 596         | AAGATGTGGCGCTTCTGAGAGCGGTCACTTTAACTTCCCTGAGCCCGCAATGAGATCACC    | 655          |       |            |     |        |        |      |
| QY | 732         | CGATGTCACACCGGAGTGAATGACATTTGTGCATTCATGTCGACATCTGAAAGCA         | 791          |       |            |     |        |        |      |
| Db | 656         | CGATGTCACACCGGAGTGTGAGGACATTTGTGCATTCATGTTGGCATCTGAAAGGA        | 715          |       |            |     |        |        |      |
| QY | 792         | CAAAATTTGCAATCTATTTCAAAAGGCAATCAATGTCATATGTCATATGAAATGAGGA      | 851          |       |            |     |        |        |      |
| Db | 716         | CAAAATTTGCAATCTATTTCAAAAGGCAATCAATGTCATATGAAATGAGGA             | 775          |       |            |     |        |        |      |
| QY | 852         | AAAAACATGGCCCTTGGGATGATCACTATTCATATTTTTCGTTCTGTGTCCTTTTAA       | 911          |       |            |     |        |        |      |
| Db | 776         | AAAAACATGGCCCTTGGGATGATCACTATTCATATTTTCTGTGTCCTTTTAA            | 835          |       |            |     |        |        |      |
| QY | 912         | TTATTAACGGCGGCAACCTGTGGCTATTTTATCTTTTATCTCTCGAAGGCTACGAAAG      | 971          |       |            |     |        |        |      |
| Db | 836         | TTATTAACGGCGCAACCGTGGCTATTTTATCTTTTATCTCTCGAAGGCTACGAAAG        | 895          |       |            |     |        |        |      |
| QY | 972         | CAGAAGCTCAAGACAGAGACAGAGGCAATTAAGCGACATCTTAAAAAGCTATTGGA        | 1031         |       |            |     |        |        |      |
| Db | 896         | CAGAAGCTCAAGACAGAGACAGAGGCAATTAAGCGACATCTTAAAAAGCTATTGGA        | 955          |       |            |     |        |        |      |
| QY | 1032        | GGCTTCAACTAGCACTTGAACCAAGGACCAAGAAATGGCCCTGATGTGAATGAT          | 1091         |       |            |     |        |        |      |
| Db | 956         | AGCTTCAACTCGACCTTGAACCAAGGACCAAGAAATGGCCCTGATGTGAATGAT          | 1015         |       |            |     |        |        |      |
| QY | 1092        | GAGCTGTGATGATGATTTGATTAACCAATGATTTTGGATGCACTTAAAGTGAAC          | 1151         |       |            |     |        |        |      |
| Db | 1016        | GAGCTGTGATGATGATTTGATTAACCAATGATTTTGGATGCACTTAAAGTGAAC          | 1075         |       |            |     |        |        |      |
| QY | 1152        | ATATTTTCATAGACATGTGTGACCAATGCTGTTAAAAACAAGTCTGCCCATCT           | 1211         |       |            |     |        |        |      |

```
/organism="Mus musculus"  
/mol_type="mrna"  
/strain="CD-1"
```





```

/mol_type="rRNA"
/strain="DBx/2"
/db_xref="taxon:10090"
/clone="11.3.7"
/cell_type="anergic T cell"
CDS
1..1287

```

| BASE COUNT | 308 a | 320 c | 359 g | 300 f |
|------------|-------|-------|-------|-------|
| ORIGIN     |       |       |       |       |

|                            |        |                     |           |              |
|----------------------------|--------|---------------------|-----------|--------------|
| Query Match                | 62.4%; | Score 1106.8;       | DB 10;    | Length 1287; |
| Best Local Similarity      | 91.2%; | Pred. No. 2.1e-245; |           |              |
| Matches 1174; Conservative | 0;     | Mismatches 113;     | Indels 0; | Gaps 0;      |

|    |     |   |     |
|----|-----|---|-----|
| Qy | 263 | ATGGGGGCGGCGCGCTGGAGGCGCGGAGGTCCTCCGCGCGCGGAGGCTGCGCGGCTTTTCCAAATGG | 322 |
| Db | 1   | ATGGGGGCGGCGCGCGCGGAGGTCATCGCGCGCGGCGCTGGAGGAGCTGCCCGGCTA           | 60  |
| Qy | 323 | CTGGGCAATGATGCTTCTCTGCGGCGCTGAGTCCGAGGCAACCCGATTCTCCGGGAGGCTGAA     | 382 |
| Db | 61  | CTGGGCTATGATGCTCTCTCTGAGTCCGAGCGCCCGGATTCTCCGGGAGGCGGAA             | 120 |
| Qy | 383 | GCAGGTGGAGACCGGATACCTCAACGTATCCCTGGGCGAGTTCCGACACAGGAGTGAACCGT      | 442 |
| Db | 121 | GCGGTATGAGACCGGATACCTCAACGTATCTGGGCGAGTTCCGACACCGAGTGAACCGC         | 180 |
| Qy | 443 | ACGGTGTGGAGCTGAGCCGAGAGGAGCGGTGAACGGCACGAGACTCGCCGCTGAGACCTGTG      | 502 |
| Db | 181 | ACGGGTGTGGAGCTGAGCCGAGAGGAGCGGTGAACGGCACCTCGCCCTGAGACCCGTC          | 240 |
| Qy | 503 | GCTGGGAGCTCTGGATCCGCGCCGAGCGGCGCGGGCGCTTAAGCCTGTGAACCCGCGACAG       | 562 |
| Db | 241 | TCCGGGAGCTCTGGATCCGCGCCGAGCGGCGCGGGCGCTGAACGCTGTGAACCCGCGACCC       | 300 |
| Qy | 563 | AATTTACAAGGTGCCACAGGTTTGGGAGAGCACCGTGAAGTCTCTGTGGTGGCCCTCATC        | 622 |
| Db | 301 | AATTTACAAGGTGCCACAGGTTTGGGAGAGCACCGTGAAGTATCTTGGTGGCCCTCATC         | 360 |
| Qy | 623 | CAACCGCGCGGAGGCTGACCTTGGGACACAAATCCATCTGGCTTATGAGATGTGGCG           | 682 |
| Db | 361 | CAACCGCGGTGAGGCTGACCTTGGCGACAAATCCATCTGGCTTATGAGAGGAGGCGT           | 420 |
| Qy | 683 | TCTGGAGCGGTCACTTTAACTTCCCGGGAGCCCGCAATAGGTATCCCATATCTCAC            | 742 |
| Db | 421 | TCTGGAGCGGTCACTTTAACTTCCCTGGAGCCCGCAATAGGTATCCCATATCTCAC            | 480 |
| Qy | 743 | CCGGGTGCAATGACATTTGTGCAATCAATGATGGGCAATGTGAAGGACACAAAATTTGTG        | 802 |
| Db | 481 | CCGGGTGCTGGGGACAATTTGTCAATCAATGATTTGGCAATGTGAAGGACACAAAATTTGTG      | 540 |
| Qy | 803 | CAATCATTTCAAAAGGCATCAAAATGACATAGTGCATAGACTGAGGAAAAAACAATGGC         | 862 |
| Db | 541 | CAGTATTTCAAAAGGCATCAAAATGACATAGTGCATAGAGTGGGAAAAAACAACATGGC         | 600 |
| Qy | 863 | CCCTGGGTGATCATATTCATATTTTTCGTTCTCTGTGCTCTTTTATATTAATTAACGGCG        | 922 |
| Db | 601 | CCCTGGGTGATCATATTCATATTTTCTTGTTTCTGTGCTCTTTTATATTAATTAACGGCA        | 660 |
| Qy | 923 | GGAATGTGGGCTATTTTATCTTTATCTTCTGTGGAAGGCTAGAGGATGCAAGAGCTCA          | 982 |
| Db | 662 | GGAATCGTGGGCTATTTTATCTTTATCTTGTGGAAGTATGGAAGGATGCAAGAGCTCA          | 720 |

|    |      |   |      |
|----|------|---|------|
| Qy | 993  | AGCAGAGACAGAGGCATTTAAAGCAGATGCTAAAAAGCTATGTGAGAGCTCAACTA      | 1042 |
| Db | 721  | AGCAGAGAGAGAGGCATTTAAAGCAGATGCTAAAAAGCTATGTGAAAGCTCAAGCTG     | 780  |
| Qy | 1043 | CGCAGACTGTAACAGAGACAGAGAAATGGCCCTGATGAGATAGTGTGTGTGTGC        | 1102 |
| Db | 761  | CGCACCTTGTAACAGAGAGACAGAGAAATGGCCCTGATGAGATAGTGTGTGTGTGC      | 840  |
| Qy | 1103 | ATTGAATGTGATTAACCAATGATTTGGTAGCATCTTAAAGTGCACCAATTTTCCAT      | 1162 |
| Db | 841  | ATTGAGCTCATTAAGCCAAATGATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGC     | 900  |
| Qy | 1163 | AAGCATGTGTGACCCATGTGCTGTTTAAACACAGACTTGGCCCATATGSCAAATGTGAC   | 1222 |
| Db | 901  | AAGCAGATGTGTGACCCGTGTGCTTTTGAACACAGACTTGGCCCATATGTCAGTGTGAC   | 960  |
| Qy | 1223 | ATATCTAAAGCTTTGGGAAATGAGGTGAGATGTTGAAGATGAGATCAGTCTTTTACAAGTC | 1282 |
| Db | 961  | ATTTCTAAAGCTCTGGGAAATGAGGTGAGATGTTGAAGATGAGATCAGTCTTTTACAAGTT | 1022 |
| Qy | 1283 | CCTGTATCCATGAAATATCTTAATAGTGTCTCTCCCATGAAGAGATATGCGACGAG      | 1344 |
| Db | 1021 | CCTGTATCTTAAGAGCATCTTAATAGTGTCTCTCCCATGAAGAGACATGTCCCAATGTAG  | 1080 |
| Qy | 1343 | ACCCATCATCTGATATGCTTACATGACAGAGAAACAGATGAACCGCTCTGAGAGAAACAC  | 1402 |
| Db | 1081 | AGTGCATCATCTGATATGCTTACATGACAGAGAGAGATGAGCCACTCTGAGAGAAAT     | 1142 |
| Qy | 1403 | GTCAGATCAACAAATGAAATGCTTACAGCTGTGTAAACATGAGCAATTTCTGTGTGACGTG | 1462 |
| Db | 1141 | GCGAGATCAACAAATGAAATATCTACAGCTGTGTAAACATGAGCAATTTCTGTGTGACGTG | 1200 |
| Qy | 1463 | GATGTATTTCTCATGTGACCAACCAACCTTTGAGAGACGAAATCTCTTAATCAAGAG     | 1522 |
| Db | 1201 | GATGTGTTCCTCATGTGTGACCAACCAACCTTTGAGAGAGATGAAATCTCTGTATCAAGAG | 1260 |
| Qy | 1523 | ACTGTCTTTGAGAAATTAATCTTAA                                     | 1549 |
| Db | 1261 | GCAAGCTTTGAGAGATTAATCTTAA                                     | 1287 |

| RESULT 9 | LOCUS    | DEFINITION | ACCESSION | VERSION    | KEYWORDS                                   | SOURCE                | ORGANISM      | REFERENCE | AUTHORS   | TITLE                             | JOURNAL     | REFERENCE           | AUTHORS  | TITLE             | JOURNAL                 | COMMENT  |   |
|----------|----------|------------|-----------|------------|--|-----------------------|---------------|-----------|---|-----------------------------------|-------------|---------------------|--|-------------------|-------------------------|--|---|
| AK027169 | AK027169 | AK027169   | AK027169  | AK027169.1 | oligo capping, fls (full insert sequence). | Human sapiens (human) | Human sapiens | 1 (sites) | Kanakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Masumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Ota, T., Suzuki, Y., Okeyashima, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isegai, T. and Sugano, S. | NEO human cDNA sequencing project | Unpublished | 2 (bases 1 to 1871) | Sugano, S., Suzuki, Y., Ota, T., Okeyashi, M., Nishi, T., Isegai, T., Shibahara, T., Tanaka, T. and Nakamura, Y. | Direct Submission | Submitted (29-JUG-2000) | Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Humanae Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cda@sums-u-tokyo.ac.jp, Tel:81-3-5445-5286, Fax:81-3-5449-5416) | NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- and 3'-end one pass sequencing; Department of |

Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).

FEATURES  
Source  
1..1871  
Location/Qualifiers

CDS

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="LNG04848"  
/issue\_type="human lung"  
/clone\_id="LNG"  
/note="Cloning vector pME18SFL3"  
76..546  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BA15682.1"  
/db\_xref="GI:10440233"  
/translation="MEIVYLCAINCINOMITWISILNCHHIFHKVCVDPWLEHRTCPM  
CKDILKALGIVDYDGSASIQVPSNHSNSASHSRSTASGTAAGTQDDE  
PPLREHVOSTNBSLOIVNBSANSVAVDVIPIHVDNPFEDFTPOQTAVALIKS"

BASE COUNT 670 a 277 c 321 g 603 t  
ORIGIN

Query Match 40.3%; Score 714.4; DB 9; Length 1871;  
Best Local Similarity 97.9%; Pred. No. 1.7e-154;  
Matches 755; Conservative 0; Mismatches 12; Indels 4; Gaps 3;

1005 AGGCAAGATGCTAAAGAACTATTGGAAGGCTTCAACTAGCGACACTGAAACAGAGAC 1064  
1 AGGCAAGATGCTAAAGAAAGCTGAGAGAGCTTCAACTAGCGACACTGAAACAGAGAC 60  
1065 AGGAAATGGCCCTGAG 1124  
61 AGGAAATGGCCCTGAG 120  
1125 ATTGGTAC-GCATCTTACGTGCAACAATTTCCATAAGACAGTGTGACCATG 1183  
121 ATTGGTACAGCATCTTACGTGCAACAATTTCCATAAGACAGTGTGACCATG 180  
1184 CTGTTAAACACAGAGCTTCCCTCCATGTCACAAATGTGACATCTCAAGCTTGGGAAT 1243  
181 CTGTTAAACACAGAGAGCTTCCCTCCATGTCACAAATGTGACATCTCAAGCTTGGGAAT 240  
1244 GAGTGAATGTTGAAGATGAGATGATGCTTTCACAGTCCCTGTATCCATGAATATCT 1303  
241 GAGTGAATGTTGAAGATGAGATGATGCTTTCACAGTCCCTGTATCCATGAATATCT 300  
1304 AATAGTCCCTCTCCATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1363  
301 AATAGTCCCTCTCCATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
1364 TCAAGTACAGGAGACAGATGAAACCGCTCCGAGAGAGACAGTGCAGTCAACAAGTAAAGT 1423  
361 TCAAGTACAGGAGACAGATGAAACCGCTCCGAGAGAGACAGTGCAGTCAACAAGTAAAGT 420  
1424 CTACAGCTGTAAACATGAGAGAAATTCGTGACAGTGTATTCCTCAGTTGAC 1483  
421 CTACAGCTGTAAACATGAGAGAAATTCGTGACAGTGTATTCCTCAGTTGAC 480  
1484 AACCAACCTTTGAAG 1543  
481 AACCAACCTTTGAAG 540  
1544 TCTTAAATCTGTGTAATGAGAACTTGAACCATAGTATGATTAACAGAGAGAGAG 1603  
541 TCTTAAATCTGTGTAATGAGAACTTGAACCATAGTATGATTAACAGAGAGAGAG 600  
1604 GGCCTAGTTCTATTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1663  
601 GGCCTAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
1664 CTCAGATGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1723

Db 661 CTCAGATGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 718  
QY 1724 TTTTCCACCAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1774  
Db 719 TTTTCCACCAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 768

RESULT 10  
AL391315  
LOCUS  
DEFINITION  
Human DNA sequence from clone RP11-150F24 on chromosome X, complete  
sequence.  
ACCESSION  
AL391315  
VERSION  
AL391315.22 GI:14575271  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 169812)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (27-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
request: clonerequest@sanger.ac.uk  
On Jun 28, 2001 this sequence version replaced gi:14455902.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Bm, EMBL; Sw,  
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome X, constructed by the Sanger Centre Chromosome X Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/ChrX  
RP11-150F24 is from the library RPCI-11.1 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP11-150F24. The true  
left end of clone RP11-321G1 is at 66861 in this sequence. The true  
right end of clone RP11-697G3 is at 161410 in this sequence.

## FEATURES

## Source

1..169812  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP11-150F24"  
/clone\_id="RPCI-11.1"  
12..569  
/note="match: GGS: Bm:AQ375615"

## misc\_feature

94..165  
/note="2 copies 36 mer 90% conserved"

## repeat\_region

512..729  
/note="MIR repeat: matches 3..238 of consensus"

## repeat\_region

1338..1638  
/note="AluY repeat: matches 1..304 of consensus"

## repeat\_region

3599..3642  
/note="12 repeat: matches 2707..2750 of consensus"

repeat\_region 3722..3899  
/note="MIR repeat: matches 12..171 of consensus"  
repeat\_region 3915..4028  
/note="MIR repeat: matches 131..238 of consensus"  
repeat\_region 5121..5197  
/note="L2 repeat: matches 2611..2710 of consensus"  
repeat\_region 6734..7029  
/note="AluSg repeat: matches 1..297 of consensus"  
repeat\_region 7660..7955  
/note="AluSg repeat: matches 1..308 of consensus"  
repeat\_region 8891..8940  
/note="L1PA9 repeat: matches 6114..6163 of consensus"  
repeat\_region 10067..10370  
/note="AluSx repeat: matches 1..297 of consensus"  
repeat\_region 10411..10695  
/note="AluSx repeat: matches 1..288 of consensus"  
repeat\_region 10737..10879  
/note="MER3 repeat: matches 5..154 of consensus"  
repeat\_region 11078..11380  
/note="AluY repeat: matches 1..300 of consensus"  
repeat\_region 11571..12560  
/note="L1MD2 repeat: matches 5328..6331 of consensus"  
repeat\_region 13476..13794  
/note="L1MS1 repeat: matches 5837..6165 of consensus"  
repeat\_region 13906..14041  
/note="L2 repeat: matches 2536..2671 of consensus"  
repeat\_region 14106..14411  
/note="AluY repeat: matches 2..306 of consensus"  
repeat\_region 16413..16704  
/note="AluSx repeat: matches 1..295 of consensus"  
repeat\_region 17566..17697  
/note="FLAM C repeat: matches 1..132 of consensus"  
repeat\_region 17839..18080  
/note="MIR repeat: matches 3..260 of consensus"  
repeat\_region 18961..19153  
/note="MIR repeat: matches 30..262 of consensus"  
repeat\_region 19562..19658  
/note="MER3A repeat: matches 12..109 of consensus"  
repeat\_region 20326..20413  
/note="L1MC4 repeat: matches 7861..7948 of consensus"  
repeat\_region 20471..20754  
/note="AluSg repeat: matches 1..286 of consensus"  
repeat\_region 20755..20782  
/note="L4 copies 2 mer ta 8% conserved"  
repeat\_region 20999..21184  
/note="L2 repeat: matches 2529..2729 of consensus"  
repeat\_region 21194..21952  
/note="L2 repeat: matches 1730..2546 of consensus"  
repeat\_region 22817..23035  
/note="MIR repeat: matches 15..238 of consensus"  
misc\_feature 23641..24552  
/note="Cpg island"  
/evidence="not\_experimental"  
repeat\_region 24009..24384  
/note="L1M3 repeat: matches 310..694 of consensus"  
misc\_feature complement(24313..24743)  
/note="match: GSS: Em:AQ23340"  
repeat\_region 25462..25711  
/note="MIR repeat: matches 2..256 of consensus"  
repeat\_region 25838..25939  
/note="MIR repeat: matches 6..106 of consensus"  
repeat\_region 26420..26566  
/note="MIR repeat: matches 2..151 of consensus"  
repeat\_region 28132..28302  
/note="MIR repeat: matches 64..238 of consensus"  
repeat\_region 29706..29835  
/note="L2 repeat: matches 2606..2746 of consensus"  
repeat\_region 30126..30458  
/note="MER2 repeat: matches 2..344 of consensus"  
repeat\_region 30873..31184  
/note="AluSx repeat: matches 1..312 of consensus"  
repeat\_region 34418..34503  
/note="L1PA3 repeat: matches 6061..6146 of consensus"

repeat\_region 34518..34851  
/note="MER2 repeat: matches 1..344 of consensus"  
repeat\_region 35308..35457  
/note="MIR repeat: matches 1..162 of consensus"  
repeat\_region 35828..36039  
/note="MIR repeat: matches 7..226 of consensus"  
repeat\_region 36274..36349  
/note="L19 copies 4 mer ttcg 76% conserved"  
repeat\_region 36353..36642  
/note="AluSx repeat: matches 1..291 of consensus"  
repeat\_region 37753..37919  
/note="FRAM repeat: matches -1..165 of consensus"  
repeat\_region 38928..39134  
/note="L2 repeat: matches 2467..2703 of consensus"  
repeat\_region 39127..39319  
/note="L2 repeat: matches 2031..2228 of consensus"  
repeat\_region 39546..39726  
/note="MER3A repeat: matches 1..186 of consensus"  
repeat\_region 39738..40910  
/note="Charliela repeat: matches 1..1258 of consensus"  
repeat\_region 40911..41435  
/note="L1PA5 repeat: matches 5616..6143 of consensus"  
repeat\_region 41436..41674  
/note="Charliela repeat: matches 1258..207 of consensus"  
repeat\_region 41675..41942  
/note="AluY repeat: matches 1..283 of consensus"  
repeat\_region 41943..42098  
/note="Charliela repeat: matches 207..1455 of consensus"  
repeat\_region 42634..42699  
/note="WSP repeat: matches 356..426 of consensus"  
repeat\_region 43609..43659  
/note="L2 repeat: matches 2696..2746 of consensus"  
repeat\_region 45090..45187  
/note="L2 repeat: matches 2606..2705 of consensus"  
repeat\_region 45672..46211  
/note="L1MA4 repeat: matches 5700..6295 of consensus"  
repeat\_region 46808..47107  
/note="Aluub repeat: matches 1..301 of consensus"  
repeat\_region 47190..47271  
/note="HY1 repeat: matches 28..111 of consensus"  
repeat\_region 47355..47455  
/note="MER3 repeat: matches 5..734 of consensus"  
repeat\_region 47525..47693  
/note="MIR repeat: matches 3..196 of consensus"  
misc\_feature 47826..47978  
/note="match: GSS: Em:AQ136922"  
misc\_feature 47985..48294  
/note="match: GSS: Em:AQ136922"  
repeat\_region 48203..50340  
/note="L1M1 repeat: matches -1389..1172 of consensus"  
repeat\_region 50335..51398  
/note="L1PA2 repeat: matches 1679..2716 of consensus"  
repeat\_region 51399..51743  
/note="Aluub repeat: matches 1..305 of consensus"  
repeat\_region 51744..52940  
/note="L1PA2 repeat: matches 2716..3982 of consensus"  
repeat\_region 52917..54922  
/note="L1PA2 repeat: matches 4138..6146 of consensus"  
repeat\_region 54936..57273  
/note="L1MA3 repeat: matches 3551..6304 of consensus"  
repeat\_region 57453..58059  
/note="L1M3 repeat: matches 5276..5875 of consensus"  
repeat\_region 58056..58804

Query Match 38.0% Score 674.4 DB 9 Length 169812;  
Best Local Similarity 99.9% Pred.No.9.1e-145 Indels 0 Gaps 0;  
Matches 675; Conservative 0; Mismatches 1;

QY 74 CGAGAGCTGATCTGGGCACTGTGTGCTGACGCTACGTCCTCTCGCTCGAC 133  
DB 23909 CGAGAGCTGATCTGGGCACTGTGTGCTGACGCTACGTCCTCTCGCTCGAC 23968  
QY 134 TAGCTCCAGACTCCCACTGCTCACTTCCTCTCCCACTGGGCGACCTGTCAAG 193

```

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hunquerry@sanger.ac.uk
Project Information
Center project name: bA697G3
Summary Statistics
Assembly program: XGAP4, version 4.5
Sequencing vector: plasmid; L08752; 100% of reads

```

|    |       |   |       |
|----|-------|---|-------|
| QY | 74    | CCGAGGAGCTGATCTCTGGGCAACTGTGTCTACAGCTACGTGCTCTGGCTCCAGAC        | 133   |
| Db | 23909 | CCGAGGAGCTGATCTCTGGGCAACTGTGTGTACGTACGTACGTGCTCTGGCTCCAGAC      | 23968 |
| QY | 134   | TACCTGGACAGCTCCCGCAGTCTACATCCATTTCTTCCCACTATGGGCGGCACTGCTTAAG   | 193   |
| Db | 23969 | TACCTGGACAGCTCCCGCAGTCTACATCCATTTCTTCCCACTATGGGCGGCACTGCTTAAG   | 24028 |
| QY | 194   | ACCAAGGCTCTGTCGAACGCTGTAAGAGGGGCGGTGTCAGAGGGGCGCTAGAGAACTGCGAGC | 253   |
| Db | 24029 | ACCAAGGCTCTGTCGAACGCTGTAAGAGGGGCGGTGTCAGAGGGGCGCTAGAGAACTGCGAGC | 24088 |
| QY | 254   | GCGCGCGCCATGGGGCCGCGCTGGGGGCGGGGTCTCTGTCGGCGGTGGGCTGGGGCTTT     | 313   |
| Db | 24089 | GCGCGCGCCATGGGGCCGCGCTGGGGGCGGGGTCTCTCTGCGCGGTGGGCTGGGGCTTT     | 24148 |
| QY | 314   | TCCAGATTGCTGGCATGTGTCCTTCTGTGTCGCTCTGATGTCGAGAGCACCGGATTCGCG    | 373   |
| Db | 24149 | TCCAGATTGCTGGCATGTGTCCTTCTGTGTCGCTCTGATGTCGAGAGCACCGGATTCGCG    | 24208 |
| QY | 374   | GGGGCTGAAGGACGTGTGGAACCGGGTACCTGCACGCTGTCTGGGGGTTTCGGCACACGGGA  | 433   |
| Db | 24209 | GGGGCTGAAGGACGTGTGGAACCGGGTACCTGCACGCTGTCTGGGGGTTTCGGCACACGGGA  | 24268 |
| QY | 434   | GTGAACCGTACCGGTGTGGAGCTGAGCCAGAGAGGGCGTGTACCGGACGAGCATGCGCGCTG  | 493   |
| Db | 24269 | GTGAACCGTACCGGTGTGGAGCTGAGCCAGAGAGGGCGTGTACCGGACGAGCATGCGCGCTG  | 24328 |
| QY | 494   | GAGCTGTGGCTGGGGGTCTCGGTACCGGCCGACGAGGCGCGGGGCGCTTAACGCGTGTAAAC  | 553   |
| Db | 24329 | GAGCTGTGGCTGGGGGTCTCGGTACCGGCCGACGAGGCGCGGGGCGCTTAACGCGTGTAAAC  | 24388 |
| QY | 554   | CCGACACGCAATTTTCACGATGCCACGATGTTTGGGAAAGCACCGTGCAGATCTTCTTGATTG | 613   |
| Db | 24389 | CCGACACGCAATTTTCACGATGCCACGATGTTTGGGAAAGCACCGTGCAGATCTTCTTGATTG | 24448 |
| QY | 614   | GCCCTCATCAACCGGGGCGGGGCTGCACCTTCCGACACAAGATTCATGTGGCTTAAGAG     | 673   |
| Db | 24449 | GCCCTCATCAACCGGGGCGGGGCTGCACCTTCCGACACAAGATTCATGTGGCTTAAGAG     | 24508 |
| QY | 674   | AAGTGGGCGTCTGAGCGCGTCAATCTTTAACTTCCCGGAAACCGCGATAGAGGTCAATCCC   | 733   |
| Db | 24509 | AAGTGGGCGTCTGAGCGCGTCAATCTTTAACTTCCCGGAAACCGCGATAGAGGTCAATCCC   | 24568 |

|            |  |                 |       |
|------------|--|-----------------|-------|
| OY         | 734  | ATGCTCACCGGGGTG | 749   |
| Db         | 24565  | ATGCTCACCGGGGTG | 24584 |
| RESULT 12  | ACI14183/c   |                 |       |
| LOCUS      | ACI14183   |                 |       |
| DEFINITION | Rattus norvegicus clone CH230-230F6, *** SEQUENCING IN PROGRESS  |                 |       |
| ACCESSION  | ACI14183   |                 |       |
| KEYWORDS   | HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.   |                 |       |
| SOURCE     | Rattus norvegicus (Norway rat).  |                 |       |
| ORGANISM   | Rattus norvegicus<br>Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;<br>Rattus.   |                 |       |
| REFERENCE  | 1 (bases 1 to 207775)  |                 |       |
| AUTHORS    | Munzy,D.Marie, Metzner,M.Lee, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angilano,D., Anyelebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benhamed,F., Biswal,K., Blair,J., Blankenburg,K., Blythe,P., Brown,M., Bryson,N., Buhay,C., Burch,P., Burrell,K., Calderon,B., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Ande,C., Dedrich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Bayes,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georgegeorgis,B., Geer,K., Gill,I., Grady,M., Guerra,W., Guervara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Hareyre,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladin,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huliyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Liang,H., Johnson,B., Johnson,R., Jolyvet,A., Karpaty,S., Kelly,S., Khan,Z., King,L., Kovac,C., Kowals,C., Kraft,C.L., Lebow,H., Levay,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louisedge,H., Lozano,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Napua,P., Martin,K., Martin,R., Martinez,B., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Ming,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankevici,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwokkeleneh,O., Okumu,G., Olanrunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Plamkoch,C., Plopper,F., Poundexter,A., Popovic,D., Primis,E., Pu,L., L., Putao,B., Quirroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodeky,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smag's,D., Snead,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinel,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Thong,A., Trejos,Z., Uমান,K., Valse,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Ward,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A. |                 |       |
| TITLE      | Unpublished Direct Submission  |                 |       |
| JOURNAL    | 2 (bases 1 to 207775)  |                 |       |
| REFERENCE  | Worley,K.C.  |                 |       |

```

REFERENCE
AUTHORS
JOURNAL
TITLE
Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 207775)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23605528.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GLPV
Center clone name: CH230-230F6

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 191020 bases at least Q40
Consensus quality: 194390 bases at least Q30
Consensus quality: 197174 bases at least Q20
Estimated insert size: 203127; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*
* 1 57140: contig of 57140 bp in length
* 57240: gap of unknown length
* 57241 195381: contig of 138141 bp in length
* 195382 195481: gap of unknown length
* 195482 196790: contig of 1309 bp in length
* 196791 196890: gap of unknown length
* 196891 207775: contig of 10885 bp in length.
Location/Qualifiers
1.207775
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-230F6"
223..1082
/note="clone_boundary
clone_end:17
site:BCOI
end_sequence:EZ112338"

BASE COUNT 61791 a 39195 c 38071 g 59202 t 9516 others
ORIGIN
Query Match 24.4% Score 432.4; DB 2; Length 207775;
Best local Similarity 90.1%; Pred. No. 1,1e-08;
Matches 463; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```

|  |                                      |
|--|--------------------------------------|
| Consensus quality: 245788 bases at least Q20   |                                      |
| Insert size: 249205; sum-of-contigs  |                                      |
| Insert size: 16861; 5.2% error; agarose-fp   |                                      |
| Quality coverage: 5.70x in Q20 bases; sum-of-contigs coverage: 11.65x in Q20 bases; agarose-fp |                                      |
| -----  |                                      |
| NOTES: This is a 'working draft' sequence. It currently  |                                      |
| * consists of 32 contigs. The true order of the pieces   |                                      |
| * is not known and their order in this sequence record is                                      |                                      |
| * arbitrary. Gaps between the contigs are represented as                                       |                                      |
| * runs of N, but the exact sizes of the gaps are unknown.                                      |                                      |
| * This record will be updated with the finished sequence                                       |                                      |
| * as soon as it is available and the accession number will                                     |                                      |
| * be preserved.  |                                      |
| 1  | 40466: contig of 40486 bp in length  |
| *  | 40487                                |
| *  | 40586: gap of 100 bp                 |
| *  | 40587                                |
| *  | 53971: contig of 1385 bp in length   |
| *  | 53972                                |
| *  | 54071: gap of 100 bp                 |
| *  | 54072                                |
| *  | 65771: contig of 11700 bp in length  |
| *  | 65772                                |
| *  | 65871: gap of 100 bp                 |
| *  | 65872                                |
| *  | 70703: contig of 4832 bp in length   |
| *  | 70704                                |
| *  | 70803: gap of 100 bp                 |
| *  | 70804                                |
| *  | 72860: contig of 2057 bp in length   |
| *  | 72861                                |
| *  | 72960: gap of 100 bp                 |
| *  | 72961                                |
| *  | 75036: contig of 2076 bp in length   |
| *  | 75137                                |
| *  | 75136: gap of 100 bp                 |
| *  | 75137                                |
| *  | 77847: contig of 2711 bp in length   |
| *  | 77848                                |
| *  | 77947: gap of 100 bp                 |
| *  | 81334: contig of 3387 bp in length   |
| *  | 81335                                |
| *  | 81435: gap of 100 bp                 |
| *  | 81435                                |
| *  | 83829: contig of 2395 bp in length   |
| *  | 83830                                |
| *  | 83929: gap of 100 bp                 |
| *  | 83930                                |
| *  | 114462: contig of 33533 bp in length |
| *  | 117463                               |
| *  | 117562: gap of 100 bp                |
| *  | 117563                               |
| *  | 188799: contig of 71237 bp in length |
| *  | 188800                               |
| *  | 188809: gap of 100 bp                |
| *  | 188900                               |
| *  | 196007: contig of 9108 bp in length  |
| *  | 196008                               |
| *  | 196107: gap of 100 bp                |
| *  | 198108                               |
| *  | 200705: contig of 2598 bp in length  |
| *  | 200706                               |
| *  | 200805: gap of 100 bp                |
| *  | 200806                               |
| *  | 202952: contig of 2147 bp in length  |
| *  | 202953                               |
| *  | 203052: gap of 100 bp                |
| *  | 203053                               |
| *  | 208517: contig of 5465 bp in length  |
| *  | 208518                               |
| *  | 208617: gap of 100 bp                |
| *  | 211137: contig of 2520 bp in length  |
| *  | 211138                               |
| *  | 211237: gap of 100 bp                |
| *  | 211238                               |
| *  | 213763: contig of 2526 bp in length  |
| *  | 213764                               |
| *  | 213766: gap of 100 bp                |
| *  | 213864                               |
| *  | 216675: contig of 2812 bp in length  |
| *  | 216775                               |
| *  | 216776                               |
| *  | 219251: contig of 2476 bp in length  |
| *  | 219252                               |
| *  | 219351: gap of 100 bp                |
| *  | 219352                               |
| *  | 221818: contig of 2467 bp in length  |
| *  | 221819                               |
| *  | 221918: gap of 100 bp                |
| *  | 221919                               |
| *  | 223941: contig of 2023 bp in length  |
| *  | 223942                               |
| *  | 224041: gap of 100 bp                |
| *  | 224042                               |
| *  | 226074: contig of 2033 bp in length  |
| *  | 226075                               |
| *  | 226175                               |
| *  | 228533: contig of 2359 bp in length  |
| *  | 228534                               |
| *  | 228633: gap of 100 bp                |
| *  | 228634                               |
| *  | 231398: contig of 2765 bp in length  |
| *  | 231399                               |
| *  | 231498: gap of 100 bp                |
| *  | 231499                               |
| *  | 234893: contig of 3395 bp in length  |
| *  | 234894                               |
| *  | 234993: gap of 100 bp                |
| *  | 234994                               |
| *  | 237520: contig of 2527 bp in length  |
| *  | 237521                               |
| *  | 237620: gap of 100 bp                |
| *  | 237621                               |
| *  | 239636: contig of 2016 bp in length  |
| *  | 239637                               |
| *  | 239736: gap of 100 bp                |
| *  | 239737                               |
| *  | 241963: contig of 2227 bp in length  |
| *  | 241964                               |
| *  | 242063: gap of 100 bp                |
| *  | 242064                               |
| *  | 244300: contig of 2237 bp in length  |
| *  | 244301                               |
| *  | 244400: gap of 100 bp                |
| *  | 244531                               |
| *  | 244533: contig of 2139 bp in length  |



```
misc_feature /note="assembly_fragment:02643  
clone_end:386  
vector_side:left"  
40587..53971  
/note="assembly_fragment:02740  
fragment_chain:1"  
54072..65771  
/note="assembly_fragment:02978  
fragment_chain:1"  
65872..70703  
/note="assembly_fragment:00688  
fragment_chain:1"  
70804..72860  
/note="assembly_fragment:02769  
fragment_chain:2"  
72961..75036  
/note="assembly_fragment:04356  
fragment_chain:2"  
75137..77847  
/note="assembly_fragment:03220  
fragment_chain:3"  
77948..81334  
/note="assembly_fragment:04098  
fragment_chain:3"  
81435..83829  
/note="assembly_fragment:00010"  
83930..117462  
/note="assembly_fragment:00572.  
117563..188799  
/note="assembly_fragment:00572.  
188900..198007  
/note="assembly_fragment:01009"  
198108..200705  
/note="assembly_fragment:01021"  
200806..202952  
/note="assembly_fragment:01407"  
203053..208517  
/note="assembly_fragment:01432  
208618..211137  
/note="assembly_fragment:02161"  
211238..213763  
/note="assembly_fragment:02187  
213864..216675  
/note="assembly_fragment:02370"  
216776..219951  
/note="assembly_fragment:02391"  
219352..221818  
/note="assembly_fragment:02767  
221919..223941  
/note="assembly_fragment:02825  
224042..226074  
/note="assembly_fragment:03074"  
226175..228533  
/note="assembly_fragment:03127  
228634..231398  
/note="assembly_fragment:03137  
231499..234893  
/note="assembly_fragment:03306  
234994..237520  
/note="assembly_fragment:03320"
```

|                       |   |  |                                 |                        |  |  |
|-----------------------|---|--|---------------------------------|------------------------|--|--|
|                       | misc_feature  | 237621..239636   | /note="assembly_fragment:03414" |                        |  |  |
|                       | misc_feature  | 239787..241963   | /note="assembly_fragment:03479" |                        |  |  |
|                       | misc_feature  | 242064..244300   | /note="assembly_fragment:03553" |                        |  |  |
|                       | misc_feature  | 244401..246539   | /note="assembly_fragment:04290" |                        |  |  |
|                       | misc_feature  | 246640..249593   | /note="assembly_fragment:04434" |                        |  |  |
|                       | misc_feature  | 249694..252305   | /note="assembly_fragment:04462" |                        |  |  |
| BASE COUNT            | 73622 a   | 49859 c  | 48058 g                         | 77640 t    3126 others |  |  |
| ORIGIN                |   |  |                                 |                        |  |  |
| Query Match           | 24.2%   | Score 429.2;   | DB 2;                           | Length 252305;         |  |  |
| Best Local Similarity | 91.4%;  | Pred. No. 6;2e-86;   |                                 |                        |  |  |
| Matches 455;          | Conservative  | 0;   | Mismatches 43;                  | Indels 0;              |  |  |
| Gaps                  | 0;  |  |                                 |                        |  |  |
| Ox                    | 252   | GCAGCGGCCCACATGGAGGCCCCCGCCTGGGGGCGGGGATCTCTTCCGACGAGTGCGAGGT 311      |                                 |                        |  |  |
| Db                    | 54543   | GAGTGGCCCCCATGGAGGCCCGCCCGCGGAAACGGGGATCTATCCCGCGCGCGCGGGA 54602       |                                 |                        |  |  |
| Ox                    | 312   | TTTTCAGATTGCTGGCAATGTCCTTCCTCTCTGCGCCCTTAAGTCCGAGACACCAGTTCC 371       |                                 |                        |  |  |
| Db                    | 54603   | CTGCGCGGCTACTGCTGCTGGTGCTTCTCTCTGCTCTGAGTCTGAGTCCGACAGCGCGCTTCCC 54662 |                                 |                        |  |  |
| Ox                    | 372   | GGGGGGGCTAACAAGTAGGAGAACGGGAACTCAACAGTCATCGAGCGGGTTCGACACGG 431        |                                 |                        |  |  |
| Db                    | 54663   | GGGAGGCGCAACCGCTGTGACCGGCTACTCAAGTGTCTGGCGGGTTCGACACCG 54722           |                                 |                        |  |  |
| Ox                    | 432   | GAGTGAAACCGTACGGTGTGGAGACTGAGACGAGAGGGCGGTGACGACGAGACTGCGCCG 491       |                                 |                        |  |  |
| Db                    | 54723   | GAGTGAAACCGCACGGTGTGGAGACTGAGCGAGAGGGCGGTGACGACGAGACTGCGCCG 54782      |                                 |                        |  |  |
| Ox                    | 492   | TGGAGCCTGTGCGTGGGGGTCTGTGTACGCGCCGAGACGGGGCGCGCTTAAAGCTGTGA 551        |                                 |                        |  |  |
| Db                    | 54793   | TGGAGCCCGTCTCCCGGGTCTCTGTGTACGCGCCGAGCGGGCGCGCTCAAGCGCTGTGA 54842      |                                 |                        |  |  |
| Ox                    | 552   | ACCGGACACAGAAATTTACAGGTGCCACAGGTTGGGGAGAACCGTGAAGTCTCTGTGT 611         |                                 |                        |  |  |
| Db                    | 54843   | ACCGGACACCAATTTACAGGTGCCACAGGTTGGGGAGAACAGGTGCAAGATCTTGGT 54902        |                                 |                        |  |  |
| Ox                    | 612   | TGGCGCTCATCCAAGCGCGCGGGGCTGCACTTTCGAGACAGATCCATCTGGCTATG 671           |                                 |                        |  |  |
| Db                    | 54903   | TGGCGCTCATCCAAGCGCGGTGAGGCTGTGCACTTTCGAGACAGATCCATCTGGCTATG 54962      |                                 |                        |  |  |
| Ox                    | 672   | AGAGATGGGGGTCTGAGACCGGTCACTTTTAACTTCCCGGGAGACCGGATGAGGTCAATCC 731      |                                 |                        |  |  |
| Db                    | 54963   | AGAGAGGGGCTTCTGAGACCGGTCACTTTTAACTTCCCTGGAGACCGGCAATGAGGTATCC 55022    |                                 |                        |  |  |
| Ox                    | 732   | CCATGTCTCAACCGGGGTG 749  |                                 |                        |  |  |
| Db                    | 55023   | CCATGTCTCAACCGGGGTG 55040  |                                 |                        |  |  |
| <hr/>                 |   |  |                                 |                        |  |  |
| RESULT 14             |   |  |                                 |                        |  |  |
| LOCUS                 | AC109686  | 233985 bp  | DNA                             | linear HTG 11-OCT-2002 |  |  |
| DEFINITION            | Rattus norvegicus clone CH230-231D21.   | *** SEQUENCING IN PROGRESS   |                                 |                        |  |  |
| VERSION               | AC109686  |  |                                 |                        |  |  |
| KEYWORDS              | HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.  |  |                                 |                        |  |  |
| SOURCE                | Rattus norvegicus (Norway rat)  |  |                                 |                        |  |  |
| ORGANISM              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.   |  |                                 |                        |  |  |
| REFERENCE             | 1 (bases 1 to 233985)   |  |                                 |                        |  |  |
| AUTHORS               | Murny,D.Marie., Merker,M.Ise., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alstbrooks,S., Amin,A., Anguitano,D., Anyalbech,V., Ayvagi,A., Ayodeji,W., Bada,B., Baden,H., |  |                                 |                        |  |  |





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 14:32:48 ; Search time 3882 Seconds

(Without alignments)  
11106.682 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctggagctccacgcgcgt.....ataggcaagttccctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0, Gapept 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estma:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_estcl:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estum:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Match | Length | ID       | Description        |
|------------|--------|-------|--------|----------|--------------------|
| 1          | 1205.2 | 67.9  | 1921   | AK018582 | AK018582 Mus muscu |
| 2          | 1202   | 67.8  | 2741   | AK004847 | AK004847 Mus muscu |
| 3          | 1003   | 56.5  | 2677   | BC036301 | BC036301 Homo sapi |
| 4          | 847.6  | 47.8  | 1201   | EX343985 | EX343985 EX343985  |

|    |       |      |      |    |          |                    |
|----|-------|------|------|----|----------|--------------------|
| 5  | 836.8 | 47.2 | 1954 | 11 | BC030951 | BC030951 Homo sapi |
| 6  | 746.4 | 42.1 | 861  | 14 | CD107148 | CD107148 AGNCOURT  |
| 7  | 680.4 | 38.4 | 1806 | 11 | BC012931 | BC012931 Homo sapi |
| 8  | 673.8 | 38.0 | 811  | 12 | B1600046 | B1600046 603251431 |
| 9  | 654.2 | 36.9 | 1000 | 12 | B1600419 | B1600419 603246458 |
| 10 | 644.2 | 36.3 | 676  | 9  | AV690200 | AV690200 AV690200  |
| 11 | 641   | 36.1 | 773  | 12 | B1601096 | B1601096 603249588 |
| 12 | 623.4 | 35.1 | 840  | 11 | AK008312 | AK008312 Mus muscu |
| 13 | 622.8 | 35.1 | 1150 | 12 | B1488609 | B1488609 603021137 |
| 14 | 622.4 | 35.1 | 906  | 10 | B1615370 | B1615370 602346005 |
| 15 | 622.2 | 35.1 | 683  | 9  | AV686169 | AV686169 AV686169  |
| 16 | 615.4 | 34.7 | 996  | 10 | BG176206 | BG176206 602333351 |
| 17 | 609.8 | 34.4 | 866  | 14 | CB195525 | CB195525 AGNCOURT  |
| 18 | 577   | 32.5 | 681  | 14 | B1220592 | B1220592 602935079 |
| 19 | 576.6 | 32.5 | 832  | 14 | B1708243 | B1708243 B1708243  |
| 20 | 568   | 32.0 | 1001 | 14 | B1705053 | B1705053 B1705053  |
| 21 | 564.4 | 31.8 | 683  | 9  | AV692456 | AV692456 AV692456  |
| 22 | 557   | 31.4 | 965  | 14 | CB235518 | CB235518 AGNCOURT  |
| 23 | 548   | 30.9 | 835  | 13 | B1601345 | B1601345 AGNCOURT  |
| 24 | 542.8 | 30.6 | 557  | 12 | BM507037 | BM507037 h24805.y  |
| 25 | 537.2 | 30.3 | 904  | 10 | BF664617 | BF664617 602117901 |
| 26 | 531.2 | 29.9 | 596  | 10 | BF231341 | BF231341 253699.BA |
| 27 | 524.2 | 29.5 | 862  | 10 | BF668554 | BF668554 602123601 |
| 28 | 520.8 | 29.4 | 756  | 14 | CB989957 | CB989957 AGNCOURT  |
| 29 | 515.6 | 29.1 | 902  | 10 | BG165006 | BG165006 602343868 |
| 30 | 513.8 | 29.0 | 841  | 10 | BF696359 | BF696359 602124911 |
| 31 | 509.8 | 28.7 | 812  | 12 | B1110345 | B1110345 602900777 |
| 32 | 502.2 | 28.3 | 872  | 10 | BF699229 | BF699229 602126878 |
| 33 | 499.4 | 28.2 | 789  | 10 | BF669635 | BF669635 602120364 |
| 34 | 496   | 28.0 | 879  | 10 | BF666897 | BF666897 602121530 |
| 35 | 492.4 | 27.8 | 759  | 10 | BG400279 | BG400279 602464672 |
| 36 | 482.4 | 27.2 | 774  | 10 | BG432146 | BG432146 602496525 |
| 37 | 481.6 | 27.1 | 841  | 10 | BF666840 | BF666840 602121368 |
| 38 | 481.6 | 27.1 | 847  | 10 | BE958515 | BE958515 601645138 |
| 39 | 475.6 | 26.8 | 854  | 12 | B1653069 | B1653069 603300779 |
| 40 | 471.8 | 26.6 | 856  | 10 | BE958087 | BE958087 601644929 |
| 41 | 471.2 | 26.6 | 717  | 13 | BX101982 | BX101982 BX101982  |
| 42 | 469   | 26.4 | 680  | 10 | BE308209 | BE308209 601095109 |
| 43 | 468.6 | 26.4 | 884  | 10 | BF542034 | BF542034 602063189 |
| 44 | 466.4 | 26.3 | 835  | 10 | BE958350 | BE958350 601644838 |
| 45 | 461.4 | 26.0 | 622  | 10 | BF137973 | BF137973 601784651 |

## ALIGNMENTS

| RESULT 1 | LOCUS    | DEFINITION   | ACCESSION | VERSION    | KEYWORDS    | SOURCE            | ORGANISM                   | REFERENCE  | AUTHORS  | TITLE                                 | JOURNAL  | MEDLINE  | PUBMED |
|----------|----------|--|-----------|------------|-------------|-------------------|----------------------------|--|--|---------------------------------------|----------|----------|--------|
| AK018582 | AK018582 | Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130001F19 product:GRALL homolog [Homo sapiens], full insert sequence. | AK018582  | AK018582.1 | GI:12858361 | HCC: CAP trapper. | Mus musculus (house mouse) | Carroll, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | Genome Res. 10 (10), 1617-1630 (2000) | 20499374 | 11042159 |        |
| AK018582 | AK018582 | Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130001F19 product:GRALL homolog [Homo sapiens], full insert sequence. | AK018582  | AK018582.1 | GI:12858361 | HCC: CAP trapper. | Mus musculus (house mouse) | Carroll, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | Genome Res. 10 (10), 1617-1630 (2000) | 20499374 | 11042159 |        |



QY 672 AGAGATGGGCGTCTGGAGCCGTCATCTTTAATCCCGGAGACCCGCAATGAGATCC 731  
 Db 598 AGAGAGGGGCTTCTGGAGCGGTCTTCTTACTTCCCTGGAGACCCGCAATGAGATCC 657  
 QY 732 CCATGCTCAACCCGAGTGCAGTGAACATTTGTCATCTGATGGCAATCTTGAAGCA 791  
 Db 658 CCATGCTCAACCCGAGTGCAGTGAACATTTGTCATCTGATGGCAATCTTGAAGCA 717  
 QY 792 CAAATATTCGACATCTTCAAGAGGCAATCAAGTGAACATGTCATGGAAGTGAAG 851  
 Db 718 CAAATATTCGACATCTTCAAGAGGCAATCAAGTGAACATGTCATGGAAGTGAAG 777  
 QY 852 AAAACATGGCCCTGGGTCATCACTATTCATTTTCTGCTGCTGCTTTTCA 911  
 Db 778 AAAACATGGCCCTGGGTCATCACTATTCATTTTCTGCTGCTGCTTTTCA 837  
 QY 912 TTATTAAGGCGGCACTGTGGCTATTTTATCTTTATCTGCTGAGGCTTGGAGT 971  
 Db 838 TTATTAAGGCGGCACTGTGGCTATTTTATCTTTATCTGCTGAGGCTTGGAGT 897  
 QY 972 CAAGAGCTCAAGAGCAAGAGCAAGCAATTAAGSAGAGTAAAAAGTATGGAA 1031  
 Db 898 CAAGAGCTCAAGAGCAAGAGCAAGCAATTAAGSAGAGTAAAAAGTATGGAA 957  
 QY 1032 GGTCTCACTAGGCACTGAAACAGAGCAAGGAATTTGGCCCTGATGAGATGTT 1091  
 Db 958 AGCTTCAGCTGCGCCTCTTGAACAGAGCAAGGAATTTGGCCCTGATGAGATGTT 1017  
 QY 1092 GTGCTGTGTCATTTGATTTGAACCAATTAATTTGATGAGATCTTAACTGCAAC 1151  
 Db 1018 GTGCTGTGTCATTTGATTTGAACCAATTAATTTGATGAGATCTTAACTGCAAC 1077  
 QY 1152 ATATTTTCCATTAAGACATGTTTGAACCATGCTGTGTTAAACAAGACATTTGCCATG 1211  
 Db 1078 ATATTTTCCATTAAGACATGTTGAGACCCGAGCTTTAACAAGACATTTGCCATG 1137  
 QY 1212 GCAATGTGCAATCAACCAAGCTTTGGAAATTAAGTGTGATGTTGAAGATGATGAT 1271  
 Db 1138 GCAATGTGCAATCAACCAAGCTTTGGAAATTAAGTGTGATGTTGAAGATGATGAT 1197  
 QY 1272 CTTTACAAGTCCCTGATCCATGCAATGAATCTAATAGTCCCTCCCTCCATGAAGAGATA 1331  
 Db 1198 CTTTACAAGTCCCTGATCCATGCAATGAATCTAATAGTCCCTCCCTCCATGAAGAGATA 1257  
 QY 1332 ATCGAGCGGAGACCGCATCTTGTGATTTGCTTCACTAAGAGGAAACATGACCCCTC 1391  
 Db 1258 GTGCGAGTGAAGCTGATCATCTGATTTGCTTCACTAAGAGGAAACATGACCCCTC 1317  
 QY 1392 TGGAGGAACAGTGCAGTCAACAAATGAAGTCTAAGCTGGTAAACATGAAGCAAT 1451  
 Db 1318 TGGAGGAACAGTGCAGTCAACAAATGAAGTCTAAGCTGGTAAACATGAAGCAAT 1377  
 QY 1452 CTGTGGCAGTGAATGTTATCTCTCATGTTGACCAACCACTTTGAAGAGACGAATCTC 1511  
 Db 1378 CTGTGGCAGTGAATGTTATCTCTCATGTTGACCAACCACTTTGAAGAGACGAATCTC 1437  
 QY 1512 CTAATCAAGAGCACTGCTGTGAGAAATTAATCTT-AAAATCTGTGTAATGAAAACT 1570  
 Db 1438 CTGATCAAGAGGCACTGCTGTGAGAAATTAATCTT-AAAATCTGTGTAATGAAAACT 1497  
 QY 1571 TGAACCATTTAG-TAATAACAGACTGCGATCAGGGCTAGATT-CTATTAATAATTTG 1628  
 Db 1498 TGAACCATTTAG-TAATAACAGACTGCGATCAGGGCTAGATT-CTATTAATAATTTG 1557  
 QY 1628 ATTAATTAATAATAAGATGATGTAAGAGGCTGAGAGCTAATTAATTAATGCTAAT 1688  
 Db 1558 GTAAA--CGTAAACAAATGATGTAAGAGGCTGAGAGCTAATTAATTAATGCTAAT 1614  
 QY 1689 GTTAAATGCTTAAATAATTTAATCTGTTAATTTTCCACCAAACTCATTAATAATA 1748  
 Db 1615 GTTAA--TGCTTAAACATTTTAACTCCGAGTACGTTT--CAACAACTCAGCAATAG 1670  
 QY 1749 TTTTCAATAGCAAGTTCTCT 1771

Db 1671 TTTTCAATAGCAAGTTCTCT 1693  
 RESULT 2  
 AK004847  
 LOCUS  
 DEFINITION  
 Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300002C13 product:SRAL1 homolog [Homo sapiens], full insert sequence.  
 ACCESSION  
 AK004847  
 VERSION  
 AK004847.1 GI:12836339  
 KEYWORDS  
 HTC; CAP trapper.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.  
 REFERENCE  
 AUTHORS  
 Carninci, P. and Hayashizaki, Y.  
 TITLE  
 High-efficiency full-length cDNA cloning  
 JOURNAL  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636  
 REFERENCE  
 AUTHORS  
 1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 2049374  
 PUBMED  
 11042159  
 REFERENCE  
 AUTHORS  
 2  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwano, M., Ohara, B., Matsuda, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexed sequencer  
 JOURNAL  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE  
 20530913  
 PUBMED  
 11076861  
 REFERENCE  
 AUTHORS  
 4  
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukumitsu, Y., Komoto, H., Adachi, T., Fukuda, S., Aizawa, K., Iwano, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stenb, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, A., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Maffromi, L., Mashima, J., Mazzarelli, J., Nombart, P., Nordone, P., Ringu, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyono, K., Wang, K. H., Wetz, C., Whitaker, C., Williams, J., Wyszynski, B., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 NATURE 409 (6821), 685-690 (2001)  
 JOURNAL  
 MEDLINE  
 21085660  
 PUBMED  
 11217851  
 REFERENCE  
 AUTHORS  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs





|            |   |             |       |                 |
|------------|---|-------------|-------|-----------------|
| RESULT 3   |   |             |       |                 |
| LOCUS      | BC036901  |             |       |                 |
| DEFINITION | BC036901 Homo sapiens, clone IMAGE:553474, mRNA.  | 2677 bp     | ERRNA | linear          |
| ACCESSION  | BC036901  |             |       | HTC 04-MAR-2003 |
| VERSION    | BC036901.1  | GI:23331105 |       |                 |
| KEYWORDS   | HTC.  |             |       |                 |
| SOURCE     | Homo sapiens (human)  |             |       |                 |
| ORGANISM   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.<br>1 (bases 1 to 2677)<br>Strausberg,R.<br>Direct Submission<br>Submitted (23-AUG-2002) National Institutes of Health, Mammalian<br>Gene Collection (MGC), Cancer Genomics Office, National Cancer<br>Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590,<br>USA<br><br>NIH-MGC Project URL: http://mgc.nci.nih.gov<br>Contact: MGC help desk<br>Email: cgapds-tr@mail.nih.gov<br>Tissue Procurement: ATCC<br>CDNA Library Preparation: Life Technologies, Inc.<br>CDNA library Arrayed by: The I.M.A.G.E Consortium (ILMIN)<br>DNA Sequencing by: National Institutes of Health Intramural<br>Sequencing Center (NISC),<br>Gaithersburg, Maryland;<br>Web site: http://www.nisc.nih.gov/<br>Contact: nisc_mgc@hgri.nih.gov<br>Akheri,N., Ayele,K., Beckstrom-Stenberg,S.M., Benjamin,B.,<br>Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,<br>Dierich,N.T., Granite,S., Gan,X., Gupta,J., Haghighi,P.,<br>Ransen,N., Ho,S.-J., Karlins,S., Kwong,P., Laito,P., Legaiepi,R.,<br>Maduro,Q.L., Mastello,C., Maskeri,B., Mascarin,S.D., McCloskey,J.C.,<br>McDowell,L., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W.,<br>Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,D.,<br>Young,A., Zhang,L.-H. and Green,E.D. |             |       |                 |
| REMARK     |   |             |       |                 |
| COMMENT    |   |             |       |                 |

[illegible]

QY 1417 TGAAGTCTACAGCTGCTGTAACCATGAGCAAAATTCGTGGCAGTGGATGTATTCTCA 1476  
 Db 1254 TGAAGTCTACAGCTGCTGTAACCATGAGCAAAATTCGTGGCAGTGGATGTATTCTCA 1313  
 QY 1477 TGTGACCAACCCACCTTGAAGAGACAACTCTATCAAGAGCTGCTGTGAGGA 1536  
 Db 1314 TGTGACCAACCCACCTTGAAGAGACAACTCTATCAAGAGCTGCTGTGAGGA 1373  
 QY 1537 AATTAAATCTTAAATCTGTGTAAATGAAACCTTGAACCAATTAATTAACAGACTGC 1596  
 Db 1374 AATTAAATCTTAAATCTGTGTAAATGAAACCTTGAACCAATTAATTAACAGACTGC 1433  
 QY 1597 CAATCAGGCGCTAGTCTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656  
 Db 1434 CAATCAGGCGCTAGTCTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1493  
 QY 1657 GAAAGTCTCAGATGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716  
 Db 1494 GAAAGTCTCAGATGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1551  
 QY 1717 GTTAATCTTTTCCACCAACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1774  
 Db 1552 GTTAATCTTTTCCACCAACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1608  
 RESULT 4  
 EX343985 1201 bp mRNA linear EST 01-MAY-2003  
 LOCUS BX343985 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1043YB06 5-PRIME, mRNA sequence.  
 ACCESSION BX343985  
 VERSION BX343985.1 GI:30309065  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 11075.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1A1011ZF04QPI  
 &cluster=11075.f. Contact : Peng Liang Email : filiang@life.com  
 URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope, sequence ID : CS1A1011ZF04QPI.  
 FEATURES  
 source  
 1..1201  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1043YB06"  
 /rissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 183 a 323 c 344 g 258 t 93 others  
 ORIGIN  
 Query Match 47.8%; Score 847.6; DB 13; Length 1201;  
 Best, Local Similarity 89.7%; Pred. No. 1.2e-165;  
 Matches 931; Conservative 46; Mismatches 52; Indels 9; Gaps 7;  
 74 CCAGAGAGCTGATCTGCGACACCTGTGTGCTGACGCTACGCTCTCTGCTCCGACG 133

Db 75 CCAGAGAGCTGATCTGCGACACCTGTGTGCTGACGCTACGCTCTCTGCTCCGACG 134  
 QY 134 TACCTGCAAGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCGCACTGCTCAAG 193  
 Db 135 TACCTGCAAGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCGCACTGCTCAAG 194  
 QY 194 ACCAGGCTCTGCAAGCTGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 253  
 Db 195 ACCAGGCTCTGCAAGCTGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 254  
 QY 254 GCGCGCGCAATGAGGCT 312  
 Db 255 GCGCGCGCAATGAGGCT 314  
 QY 313 TTCCCAATTCCTGAG 372  
 Db 315 TTCCCAATTCCTGAG 374  
 QY 373 GAGGAGCTGAG 432  
 Db 375 GAGGAGCTGAG 434  
 QY 433 AGTGAACCGTACGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
 Db 435 AGTGAACCGTACGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494  
 QY 493 GAGAGCTGTGAG 552  
 Db 495 GAGAGCTGTGAG 554  
 QY 553 CCGGCAACGAATTTCAAGGTGCGCCAGGTTTGGGAAAGCAACGTCGACCTCTGTT 612  
 Db 555 CCGGCAACGAATTTCAAGGTGCGCCAGGTTTGGGAAAGCAACGTCGACCTCTGTT 614  
 QY 613 GGCCTCATCAACGCTGAG 671  
 Db 615 GGCCTCATCAACGCTGAG 674  
 QY 672 AGAGATGAGCGCTGTGAG 731  
 Db 675 AGAGATGAGCGCTGTGAG 734  
 QY 732 CCAATGCTCAACCGCTGAG 791  
 Db 735 CCAATGCTCAACCGCTGAG 793  
 QY 792 CAAAATTTGCAATCTAATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 851  
 Db 794 CAAAATTTGCAATCTAATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852  
 QY 852 AAAAAGATGAGCTGAG 911  
 Db 853 AAAAAGATGAGCTGAG 912  
 QY 912 TTATTAAGCGCGCACTGTGCGCTAATTTATCTTTAT - CTGCTCAAGAGCTACGAGAT 970  
 Db 913 TTATTAAGCGCGCACTGTGCGCTAATTTATCTTTATBTGCTGCGAGGATGAGAT 972  
 QY 971 GCAAGAGCTCAAG 1030  
 Db 973 GCAAGAGCTCAAG 1031  
 QY 1031 AGGCTTCAACTGACGACCACTGAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090  
 Db 1032 AGGCTTCAACTGACGACCACTGAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1088  
 QY 1091 TGTGCTGTGCTGCTGGA 1108  
 Db 1089 GATKAGCTTGAATTTTAA 1106  
 RESULT 5

BC030951  
 LOCUS BC030951 1954 bp mRNA linear HTC 13-JUN-2002  
 DEFINITION Homo sapiens, hypothetical protein FLJ23516, clone IMAGE:4280352,  
 mRNA.  
 ACCESSION BC030951  
 VERSION BC030951.1 GI:21410659  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 1954)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CNA Library Preparation: CLONTECH Laboratories, Inc.  
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [medpaxil.stanford.edu](mailto:medpaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 41 Row: C Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 13375696  
 This clone has the following problem: frame shifted.  
 FEATURES  
 source  
 1..1954  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="locusid:79589"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4280352"  
 /rissue\_type="Brain, Primitive neuroectodermal"  
 /clone\_lib="NIH\_MGC\_56"  
 /lab\_host="DH10B"  
 /note="vector: pDNR-LIB"  
 BASE COUNT 676 a 294 c 346 g 638 t  
 ORIGIN  
 Query Match 47.2%; Score 836.8; DB 11; Length 1954;  
 Best Local Similarity 98.7%; Pred. No. 2.2e-163;  
 Matches 864; Conservative 0; Mismatches 8; Indels 3; Gaps 2;  
 QY 900 TGTCTTTTATTTATTAAGGCGCACTGTGGCTATTTATCTTTATCTGTGCGAA 959  
 Db 1 TGTCTTTTATTTATTAAGGCGCACTGTGGCTATTTATCTTTATCTGTGCGAA 60  
 QY 960 GGCTACGAGATGCAAGAGCTCAAGAGCGAGAGCAAGCATTAAGCGAGATGCTAA 1019  
 Db 61 GGCTACGAGATGCAAGAGCTCAAGAGCGAGAGCAAGCATTAAGCGAGATGCTAA 120  
 QY 1020 AAGCTATTGGAAGGCTTCACTACGACACCTGAAACAAGAGACAGAAATTGGCCCTG 1079  
 Db 121 AAGCTATTGGAAGGCTTCACTACGACACCTGAAACAAGAGACAGAAATTGGCCCTG 180  
 QY 1080 ATGAGATAGT 1139  
 Db 181 ATGAGATAGT 240  
 QY 1140 TAACTGCAACATATTTCCATAAGACATGTGTGACCCATGGCTGTAAACACAGAA 1199

Db 241 TAACTGCAACATATTTCCATAAGACATGTGTGACCCATGGCTGTAAACACAGAA 300  
 QY 1200 CTTCGCCCATGTGCATATGTGACATCTCTCAAGCTTTGGGATTTAGGTGTGAAG 1259  
 Db 301 CTTCGCCCATGTGCATATGTGACATCTCTCAAGCTTTGGGATTTAGGTGTGAAG 360  
 QY 1260 ATGATCACTGTCTTTCAAGTCCCTGTATCCAAATATATATATATATATATATAT 1319  
 Db 361 ATGATCACTGTCTTTCAAGTCCCTGTATCCAAATATATATATATATATATATAT 420  
 QY 1320 ATGAAGAGATATATGAG 1379  
 Db 421 ATGAAGAGATATATGAG 480  
 QY 1380 ATGAACGCTCTGTGAG 1439  
 Db 481 ATGAACGCTCTGTGAG 540  
 QY 1440 ATGAACGCTCTGTGAG 1499  
 Db 541 ATGAACGCTCTGTGAG 600  
 QY 1500 ATGAACGCTCTGTGAG 1559  
 Db 601 ATGAACGCTCTGTGAG 660  
 QY 1560 ATGAACGCTCTGTGAG 1619  
 Db 661 ATGAACGCTCTGTGAG 720  
 QY 1620 ATGAACGCTCTGTGAG 1679  
 Db 721 ATGAACGCTCTGTGAG 780  
 QY 1680 ATGAACGCTCTGTGAG 1739  
 Db 781 ATGAACGCTCTGTGAG 837  
 QY 1740 ATGAACGCTCTGTGAG 1794  
 Db 838 ATGAACGCTCTGTGAG 872  
 RESULT 6  
 CD107148 861 bp mRNA linear EST 15-MAY-2003  
 LOCUS CD107148  
 DEFINITION AGSCOURT 14019634 NIH\_MGC\_179 Homo sapiens cDNA clone  
 IMAGE:3036694 5', mRNA sequence.  
 ACCESSION CD107148  
 VERSION CD107148.1 GI:30760322  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 861)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Michael Brownstein  
 CNA Library Preparation: Invitrogen Corp  
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.lnl.gov>  
 Plate: NDA426 row: k column: 23  
 High quality sequence stop: 739.  
 Location/Qualifiers  
 1..861  
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30366594"
/tissue_type="Pituitary"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances )"
/clone_1b="NH_MGC_179"
/note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
(destroyed); Site_2: NotI; Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon cloning
). Average insert size 1.1 kb. Library was constructed by
(invitrogen). Note: this is a NH_MGC Library."

```

|    | Query Match           | 42.1%  | Score 746.4 ; DB 14 ; Length 861 ; |
|----|-----------------------|--|------------------------------------|
|    | Best Local Similarity | 96.4% ;  | Ped. No. 1.3e-144 ;                |
|    | Matches 785 ;         | Conservative 0 ; Mismatches 24 ; Indels 5 ; Gaps 2 ;           |                                    |
| Qy | 963                   | TACGAAATGCAAGGCTCCAAACACAGAAACAGAGGCAATTAAAGCGAATGTCTAAAAAG    | 1022                               |
| Dy | 1                     | TACGAAATGCAAGGCTCCAAACAGAAAGAGAGGCAATTAAAGCGAATGTCTAAAAAG      | 60                                 |
| Qy | 1023                  | CTATTGSAAGGCTTCAACTACGCAACACTGAAACAAGAGACAAGSAAATTTGGCCCTGAG   | 1082                               |
| Dy | 61                    | CTATTGSAAGGCTTCAACTACGCAACACTGAAACAAGAGACAAGSAAATTTGGCCCTGAG   | 120                                |
| Qy | 1083                  | GAAATGTTGGTCGTGTGTGTGTGAATTGAATTGGTAAACAAAGATTTGGTGTGCAATCTTAA | 1142                               |
| Dy | 121                   | GAAATGTTGGTCGTGTGTGTGTGAATTGAATTGAATCAAAAGATTTGGTGTGCAATCTTAA  | 180                                |
| Qy | 1143                  | CGTGCAACCAATTTTCCATTAAGACATGTGTACCCATGGCTGTAAACAACAAGACTT      | 1200                               |
| Dy | 181                   | CGTGCAACCAATTTTCCATTAAGACATGTGTACCCATGGCTGTAAACAACAAGACTT      | 240                                |
| Qy | 1203                  | GGCCCAATGTGCAAAATGTGCAATCTCAAAAGCTTTGGGAAATTGAAGTGAATGTGAAGA   | 1262                               |
| Dy | 241                   | GGCCCAATGTGCAAAATGTGCAATCTCAAAAGCTTTGGGAAATGAGTGAATGTGAAGA     | 300                                |
| Qy | 1263                  | GATCAGTGTCTTTCAAAAGTCCCTGTATCCATGAATATCTTAATGTGGCTTCTCCAG      | 1322                               |
| Dy | 301                   | GATCAGTGTCTTTCAAAAGTCCCTGTATCCATGAATATCTTAATGTGGCTTCTCCAG      | 360                                |
| Qy | 1323                  | AAGAGATTAATGCGAGGSAACCGGATATCTGSAATAGTCTCAATGACGAGAAACAGAG     | 1382                               |
| Dy | 361                   | AAGAGATTAATGCGAGGSAACCGGATATCTGSAATAGTCTCAATGACGAGAAACAGAG     | 420                                |
| Qy | 1383                  | AACGGCTCTGGAGAAACAAGTGCAGTCAACAATGAAGCTCAAGCTGTAAACAGT         | 1442                               |
| Dy | 421                   | AACGGCTCTGGAGAAACAAGTGCAGTCAACAATGAAGCTCAAGCTGTAAACAGT         | 480                                |
| Qy | 1443                  | AAGCAAAATCTGTGCAAGTGAATGATTCCTCATGTGACAAACCAACCTTGAAGAG        | 1502                               |
| Dy | 481                   | AAGCAAAATCTGTGCAAGTGAATGATTCCTCATGTGACAAACCAACCTTGAAGAG        | 540                                |
| Qy | 1503                  | ACGAAATCTCTATACAGACATGCTGTGGAGAAATTAATCTTAAATCTGTGTAAAT        | 1562                               |
| Dy | 541                   | ACGAAATCTCTATACAGACATGCTGTGTGGAGAAATTAATCTTAAATCTGTGTAAAT      | 600                                |
| Qy | 1563                  | AGAAATCTGACCACTAGATTAATACAACTGCCATGAGGCGCTAGTTCTATTAAT         | 1622                               |
| Dy | 601                   | AGAAATCTGACCACTAGATTAATACAACTGCCATGAGGCGCTAGTTCTATTAAT         | 660                                |
| Qy | 1623                  | AATTTGATTAATTTAATTAATTAAGATGATCTGAAGAGTGTCAATGACTAATATTAAT     | 1682                               |
| Dy | 661                   | AATTTGATTAATTTAATTAATTAAGATGATGATGAAATGCTCAGATGACTAATATTAAT    | 720                                |
| Qy | 1683                  | GCTATAGTAAATAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT     | 1742                               |
| Dy | 721                   | GCTATAGT - AATAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT   | 779                                |
| Qy | 1743                  | -----ATTAATTTTCAATAGGCAAGTTCCCTTC 1772                         |                                    |
| Dy | 780                   | TATATATATTTTCAATAGGCAAGTTCCCTTC 813                            |                                    |

|            |  |       |          |         |      |        |                 |
|------------|--|-------|----------|---------|------|--------|-----------------|
| RESULT 7   | BC012931   | LOCUS | BC012931 | 1806 bp | mRNA | linear | HTC 04-MAR-2003 |
| DEFINITION | Homo sapiens, Similar to hypothetical protein FLJ23516, clone IMAGE:4456033, mRNA.   |       |          |         |      |        |                 |
| ACCESSION  | BC012931   |       |          |         |      |        |                 |
| VERSION    | BC012931.1   |       |          |         |      |        |                 |
| KEYWORDS   | HTC.   |       |          |         |      |        |                 |
| SOURCE     | Homo sapiens (human)   |       |          |         |      |        |                 |
| ORGANISM   | Homo sapiens   |       |          |         |      |        |                 |
| REFERENCE  | Bakayeta, Metzazo; Chorodeta; Craniata; Vertebrata; Euteleostomi; Mammalia; Butcheria; Primates; Catarrhini; Homnidae; Homo. Strausberg, R.  |       |          |         |      |        |                 |
| AUTHORS    | 1 (bases 1 to 1806)  |       |          |         |      |        |                 |
| TITLE      | Direct Submission  |       |          |         |      |        |                 |
| JOURNAL    | Submitted (20-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA |       |          |         |      |        |                 |
| REMARK     | NHL-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>   |       |          |         |      |        |                 |
| COMMENT    | Contact: MGC help desk   |       |          |         |      |        |                 |

CDNA Library Preparation: Life Technologies, Inc.  
 cDNA library Arranged by: The I.M.A.-E. Consortium (ILNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Guinardone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisseged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAST at: <http://image.llnl.gov>  
Series: IRAX Plate: 28 Row: b Column: 9  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21362047  
This clone has the following problem: retained intron.

```

FEATURES
source
    Location/Qualifiers
    1..1806
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4456033"
        /tissue_type="Kidney, hypernephroma"
        /clone_1b="NH MGC_89"
        /lab_host="DH10B"
        /note="vector: pCMV-Sport6"
BASE COUNT
    616 a 266 c 317 g 607 t
ORIGIN

```

|    | Query Match           | 38.44%  | Score 680.4    | DB 11    | Length 1806 |
|----|-----------------------|---|----------------|----------|-------------|
|    | Best Local Similarity | 98.6%   | Pred. 6.9e-131 |          |             |
|    | Matches 707           | Conservative 0  | Mismatches 7   | Indels 3 | Gaps 2      |
| QY | 1058                  | GGAGAAAGAGAAATGGCCCTGATGAGAAATGTTGCTGTGAGCAATGGAATGTGAATAAA | 111            |          |             |
| Db | 1                     | GGAGACAGAGAAATGGCCCTGATGAGAAATGTTGCTGTGCACTTAATGTGAATAAA    | 60             |          |             |
| QY | 1118                  | CCAAATGATTGGTAGCAATCTTAAACGTGCAACCAATATTTCCATAAGACATGTGTGAC | 117            |          |             |
| Db | 61                    | CCAAATGATTGGTAGCAATCTTAAACGTGCAACCAATATTTCCATAAGCAAGTGTGAC  | 120            |          |             |
| QY | 1178                  | CCATGCTGTGTAAACACAAGACTGGCCCATGTGCAATGTGACATATCTCAAGACTTGG  | 123            |          |             |
| Db | 121                   | CCATGCTGTGTAAACACAAGACTTGGCCCATGTGCAATGTGACATATCTCAAGACTTGG | 180            |          |             |
| QY | 1238                  | GAATTAGCTGATGTTGAGATGATCACTGTCTTTACAGTCCCTGATCCCAATGAA      | 129            |          |             |

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5303233"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_id="N1H_MGC_36"
/note="Organ. Brain; Vector: pBluescript (modified

```

|                           |        |                     |           |             |
|---------------------------|--------|---------------------|-----------|-------------|
| Query Match               | 38.0%; | Score 673.8;        | DB 12;    | Length 811; |
| Best Local Similarity     | 94.2%; | Pred. No. 1.6e-129; |           |             |
| Matches 745; Conservative | 0;     | Mismatches 37;      | Indels 9; | Gaps 4;     |

|    |     |   |     |
|----|-----|---|-----|
| QY | 74  | CCGAGAGCTGCATCTGGGGAACTTGTGTGTACGCTACGTGTCTCTGGCTCCGACG         | 133 |
| Db | 22  | CCGAGAGCTGCATCTGGGGAACTTGTGTGTACGCTACGTGTCTCTGGCTCCGACG         | 81  |
| QY | 134 | TAGCTCCAGACTCCCGAGCTCACTGCATCTCTTCCCACTGGCGGCACTGTCCAAG         | 193 |
| Db | 82  | TAGCTCCAGACTCCCGAGCTCACTGCATCTCTTCCCACTGGCGGCACTGTCCAAG         | 141 |
| QY | 194 | ACCAAGGTTCTTCGCAAGCGCTTAGAGGGCGCGTGCAGGGGCGCTAGGAACTGCCAGC      | 253 |
| Db | 142 | ACCA-GGTCTTCGCAAGCGCTTAGAGGGCGCGTGCAGGGGCGCTAGGAACTGCCAGC       | 200 |
| QY | 254 | GGCGCGCGCAATGGGG-CCGCGCTTAGGGACCGGGGTTCTCTGCACGGGTGCTCGGCTTT    | 313 |
| Db | 201 | GGCGCGCGCAATGGGGCGCGCGCTTAGGGGCTTCTCTGCACGGGTGCTCGGCTTT         | 260 |
| QY | 314 | TCCAGATTCTGGCAATGTGTGTCTCTGTCTGCTCCCTTGAATCCGACAGCACTCCGTTCCCG  | 373 |
| Db | 261 | TCCAGATTCTGGCAATGTGTGTCTCTGTCTGCTCCCTTGAATCCGACAGCACTCCGTTCCCG  | 320 |
| QY | 374 | GGGCGCTGAACAATGTGTGACCGCGTAACCTCAAGTGTCTCTGGCGGTTCCGCAAGGGAA    | 433 |
| Db | 321 | GGGCGCTGAACAATGTGTGACCGCGTAACCTCAAGTGTCTCTGGCGGTTCCGCAAGGGAA    | 380 |
| QY | 434 | GTGAACCTGTACGTTGTGGGAGCTGACGAGAGAGGGGTGTACGCGAGAGCTGGCCGCTG     | 493 |
| Db | 381 | GTGAACCTGTACGTTGTGGGAGCTGACGAGAGAGGGGTGTACGCGAGAGCTGGCCGCTG     | 440 |
| QY | 494 | GAGCTGTGGCTGGGGTCTGTGTACCGCGCAGCGGGCCCGGGGCGCTTAAAGCGCTTAAAC    | 553 |
| Db | 441 | GAGCTGTGGGTGGGGTCTGTGTACCGCGCAGCGGGCCCGGGGCGCTTAAAGCGCTTAAAC    | 500 |
| QY | 554 | CCGCAACAGAAATTTCACAGGTGGCCACGTTTTGGGAAAGACACGTGCAAGTCTGTGGTTG   | 613 |
| Db | 501 | CCGCAACAGAAATTTCACAGGTGGCCACGAAATTTGGGAAAGACACGTGCAAGTCTGTGGTTG | 560 |
| QY | 614 | GGCCTCAATCCAACGCGCGGGGGGCTGCACCTTGGCAAGACAAGATCATCTGGCTTAGAG    | 673 |
| Db | 561 | GGCCTCAATCCAACGCGCGGGGGGCTGCACCTTGGCAAGACAAGATCATCTGGCTTAGAG    | 620 |
| QY | 674 | AATAGGGGTGTGAGACCGGTACCTTTAA--CTTCCCGGGAGCCCGCAATGAG---GATC     | 727 |
| Db | 621 | AATAGGGGTGTGAGACCGGTACCTTTTAACTGTCCCGGGAGCCCGCAATGAGATCAT       | 680 |
| QY | 728 | ATCCCATGTCTCAACCCGGGTGCAATGACAAATTGTTCATC-ATGATCGGCAATCTGA      | 785 |
| Db | 681 | CCCCATGTCTCAACCCGGGTGCAATGACAAATGTTCATCATGAAATCGGCAATCTG        | 740 |
| QY | 786 | AAGGCAAAATTTCTGCAATCTATTCAAGAGGCAATCAATGACATGTGCTAGAG           | 845 |
| Db | 741 | AAGGCAAAATTTCTGCAATCTATTGCAAGAGGCAATCAATGACATGTGCTAAACAA        | 800 |
| QY | 846 | TAGGGAATAAAA 856  |     |
| Db | 801 | GTAGGGAATAA 811   |     |

BI600419 1000 bp. mRNA linear EST 07-SEP-2001  
 LOCUS 603246458P1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5288774 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI600419  
 VERSION BI600419.1 GI:15493358  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1000)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cga@dbp-remail.nih.gov  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Inceye Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: L1M11730 row: c column: 15  
 High quality sequence stop: 701.

FEATURES  
 Location/Qualifiers  
 1..1000  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5288774"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DBP103"  
 /clone\_lib="NIH\_MGC\_96"  
 /note="Organ: Brain; Vector: pBluescript (modified  
 pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3',  
 size-selected for average insert size 2.3 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 206 a 299 c 321 g 174 t  
 ORIGIN

Query Match 36.9%; Score 654.2; DB 12; Length 1000;  
 Best Local Similarity 93.7%; Pired. No. 1.9e-125;  
 Matches 747; Conservative 0; Mismatches 43; Indels 7; Gaps 6;

QY 74 CCGAGAGAGCTGATCTGCGCAACCTGTGTGTGTAAGCTACGCTCTCTGCTCCAGC 133  
 DB 22 CCGAGAGAGCTGATCTGCGCAACCTGTGTGTGTAAGCTACGCTCTCTGCTCCAGC 81  
 QY 134 TACCTGCACTCTCCCACTGCACTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 123  
 DB 82 TACCTGCACTCTCCCACTGCACTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 141  
 QY 194 ACCAGAGTCTGCAAGGCTGAGAGAGGCGCTGCAAGGAGGCTGAGAGGAGGAGC 253  
 DB 142 ACCA-GGTCTGCAAGGCTGAGAGAGGCGCTGCAAGGAGGCTGAGAGGAGGAGC 199  
 QY 254 GCGCGGCGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCT 313  
 DB 200 GCGCGGCGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCT 259  
 QY 314 TCCAGATTCTGCAAGGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCGCT 373  
 DB 260 TCCAGATTCTGCAAGGCTGAGAGGCGCGCTGAGAGGCGCGCTGAGAGGCGCGCT 319  
 QY 374 GGGGCTGAAGAGCTGTGAGCCGCTGACTCAAGCTGTCTGAGGAGGCTCCGCAACGAG 433

DB 320 GGGGCTGAAGAGCTGTGAGCCGCTGACTCAAGCTGTCTGAGGAGGCTCCGCAACGAG 379  
 QY 434 GTGAACCGTACGATGTGAGAGCTGAGCCAGAGAGGCGCTGTACCGCCAGACTGCCCTG 493  
 DB 380 GTGAACCGTACGATGTGAGAGCTGAGCCAGAGAGGCGCTGTACCGCCAGACTGCCCTG 439  
 QY 494 GAGCTGTGAGGCTGAGGCTGCTGTGTACCGCCAGAGAGGCGCGCTGTACCGCTGTAC 553  
 DB 440 GAGCTGTGAGGCTGAGGCTGCTGTGTACCGCCAGAGAGGCGCGCTGTACCGCTGTAC 499  
 QY 554 CCGCACAAGATTTCACGATGCTGAGCCAGAGGCTGAGAGGAGGCTGAGAGGCTGAGG 612  
 DB 500 CCGCACAAGATTTCACGATGCTGAGCCAGAGAGGCTGAGAGGAGGCTGAGAGGCTGAG 559  
 QY 613 GGCCTCTATCCAGACGCGCGCGCGCGCTGCG-ACCTTCCAGACAGATCCATCTGCTTATG 671  
 DB 560 GGCCTCTATCCAGACGCGCGCGCGCGCTGCG-ACCTTCCAGACAGATCCATCTGCTTATG 619  
 QY 672 AGAGATGGGCGCTGAGAGCGCTGATCTTAACTTCCCGGAGACCCGCAATGAGTCAATCC 731  
 DB 620 AGAGATGGGCGCTGAGAGCGCTGATCTTAACTTCCCGGAGACCCGCAATGAGTCAATCC 678  
 QY 732 CATGTCTGACCGCGCTGCGATGACATCTGTTGCAATCATGATCGGCAATCTGAAGGCA 791  
 DB 679 CCATGTCTGACCGCGCTGCGATGACATCTGTTGCAATCATGATCGGCAATCTGAAGG 738  
 QY 792 CAAATTTCTGCAATCTTCAAGAGAGGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 851  
 DB 739 CAAATTTCTGCAATCTTCAAGAGAGGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 796  
 QY 852 AAAACATGGCCCTTGG 868  
 DB 797 ACAAATATGGATCTTCAAGAGAGGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 813

RESULT 10  
 AV690200 676 bp mRNA linear EST 16-JAN-2002  
 LOCUS AV690200  
 DEFINITION AV690200 GRC Homo sapiens cDNA clone GRCAX02 5', mRNA sequence.  
 ACCESSION AV690200  
 VERSION AV690200.1 GI:10292063  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 676)  
 Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,  
 Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,  
 Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,  
 Hu, G., Gu, J., Chen, Z., and Han, Z.  
 Insight into hepatocellular carcinogenesis at transcriptome level  
 by comparing gene expression profiles of hepatocellular carcinoma  
 with those of corresponding noncancerous liver  
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
 21625106  
 MEDLINE  
 JOURNAL  
 PUMED 11752456  
 COMMENT Contact: Zenguan Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.

FEATURES  
 source  
 1..676  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="GRCAX02"  
 /tissue\_type="hepatocellular carcinoma"

```

/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GXC"
/notes="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      94 a      223 c      231 g      125 t      3 others
ORIGIN
Query Match      36.3%; Score 644.2; DB 9; Length 676;
Best Local Similarity 98.1%; Pred. No. 2.2e-123;
Matches 661; Conservative 0; Mismatches 11; Indels 2; Gaps 1;

QY 74 CCGAGAGCTGCACTTGCAGCAACTGTGTCTGACGCTACGTCTCTGCTCCGAG 133
Db 4 CCGAGAGCTGCACTTGCAGCAACTGTGTCTGACGCTACGTCTCTGCTCCGAG 63

QY 134 TAGCTCGAGAGCTCCCAAGTCTCACTTCCCTCCCACTGAGGAGGAGCTGCTCAAG 193
Db 64 TAGCTCGAGAGCTCCCAAGTCTCACTTCCCTCCCACTGAGGAGGAGCTGCTCAAG 123

QY 194 ACCAGAGTCTCCCAAGCGCTAGAGAGGAGCGGTGCAAGGAGGAGCTAGGAGAACTGCGAGAGC 253
Db 124 ACCAGAGTCTCCCAAGCGCTAGAGAGGAGCGGTGCAAGGAGGAGCTAGGAGAACTGCGAGAGC 183

QY 254 GCGCGCGCATGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 313
Db 184 GCGCGCGCATGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 243

QY 314 TCCAGATTGCTGAGCAAGTGTCTCTGCTGAGCTCCGAGGAGCAACCGGTTCCCGG 373
Db 244 TCCAGATTGCTGAGCAAGTGTCTCTGCTGAGCTCCGAGGAGCAACCGGTTCCCGG 303

QY 374 GGGGCTGAGAGAGTGTGAGACCGGCTACCTCACTGCTGCTGCGGAGTTCCGACACGAGGA 433
Db 304 GGGGCTGAGAGAGTGTGAGACCGGCTACCTCACTGCTGCTGCGGAGTTCCGACACGAGGA 363

QY 434 GTGAACCGTACGAGTGTGAGAGCGAGAGAGGAGCGGTGACGAGGAGCAAGTCCGCGCTG 493
Db 364 GTGAACCGTACGAGTGTGAGAGCGAGAGAGGAGCGGTGACGAGGAGCAAGTCCGCGCTG 423

QY 494 GAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 553
Db 424 GAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 483

QY 554 CCGACAGAGATTTCCAGGTGCGCCAGGTTGGGAGAGACCGGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 613
Db 484 CCGACAGAGATTTCCAGGTGCGCCAGGTTGGGAGAGACCGGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 543

QY 614 GCCCTCATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 673
Db 544 GCCCTCATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603

QY 674 AGATGAGGCTGTGAGAGCGGCTCATCTTTAACTTCCCGGAGACCGGCAATGAGTCAATCCCG 733
Db 604 AGATGAGGCTGTGAGAGCGGCTCATCTTTAACTTCCCGGAGACCGGCAATGAGTCAATCCCG 661

QY 734 ATGTCTCAACCGGAG 747
Db 662 ATGTCTCAACCGGAG 675

RESULT 11
BI601096      773 bp      mRNA      linear      EST 07-SEP-2001
LOCUS      603249588BFI NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301380 5'
DEFINITION      mRNA sequence.
ACCESSION      BI601096
VERSION      BI601096.1 GI:15494035
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE      1 (bases 1 to 773)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Miklos Balkeovits, M.D., Ph.D.
            CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LMNL at:
            http://image.lnl.gov
            Plate: LMNL1762 row: p column: 21
            High quality sequence stop: 759.
FEATURES
source
    1..773
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5301380"
    /tissue_type="hypothalamus"
    /lab_host="DH10B"
    /clone_lib="NIH_MGC_96"
    /notes="Organ: brain; Vector: pbluescriptR (modified
    pbluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (GTCGAG
    ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
    size-selected for average insert size 2.3 kb and
    normalized to 10^5. This is a primary library enriched
    for full-length clones and constructed using the
    Cap-trapper method (Carninci, in preparation) Library
    constructed by M. Brownstein (NIH/NHGRI, National
    Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT      130 a      242 c      254 g      146 t      1 others
ORIGIN
Query Match      36.1%; Score 641; DB 12; Length 773;
Best Local Similarity 96.3%; Pred. No. 1e-122;
Matches 720; Conservative 0; Mismatches 21; Indels 7; Gaps 6;

QY 74 CCGAGAGAGTGCATGTGAGGCAACTGTGTGCTGACGCTACGTCTCTGAGCTCCGAGC 133
Db 22 CCGAGAGAGTGCATGTGAGGCAACTGTGTGCTGACGCTACGTCTCTGAGCTCCGAGC 81

QY 134 TAGCTGCAAGTCCCGAGCTCACTCACTTCTCCCACTGAGCGGAGCAACCTGCTCAAG 193
Db 82 TAGCTGCAAGTCCCGAGCTCACTCACTTCTCCCACTGAGCGGAGCAACCTGCTCAAG 141

QY 194 ACCAGAGTCTCCCAAGCGCTAGAGAGGAGCGGTGCAAGGAGGAGCTAGGAGAACTGCGAGAGC 253
Db 142 ACCAGAGTCTCCCAAGCGCTAGAGAGGAGCGGTGCAAGGAGGAGCTAGGAGAACTGCGAGAGC 200

QY 254 GCGCGCGCATGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 313
Db 201 GCGCGCGCATGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 260

QY 314 TCCAGATTGCTGAGCAAGTGTCTCTGCTGAGCTCCGAGGAGCAACCGGTTCCCGG 373
Db 261 TCCAGATTGCTGAGCAAGTGTCTCTGCTGAGCTCCGAGGAGCAACCGGTTCCCGG 320

QY 374 GGGGCTGAGAGAGTGTGAGACCGGCTCAACGTTCTTGGCGGAGTTCCGACACAGGGA 433
Db 321 GGGGCTGAGAGAGTGTGAGACCGGCTCAACGTTCTTGGCGGAGTTCCGACACAGGGA 380

QY 434 GTGAACCGTACGAGTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 493
Db 381 GTGAACCGTACGAGTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 440

QY 494 GAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 553
Db 441 GAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 500

```







|    | Matches | 692; | Conservative  | 0; | Mismatches | 22; | Indels | 7; | Gaps | 5;   |
|----|---------|------|---|----|------------|-----|--------|----|------|------|
| QY | 1058    |      | GGAGACACAGGAATTGGCCCTGGATGAGAGATGGTGTGCTGTGTCATGTAATTGATATAA  |    |            |     |        |    |      | 1117 |
| Db | 1       |      | GGAGACACAGGAATTGGCCCTGGATGAGAGATGGTGTGCTGTGTCATGTA--GTAATAAA  |    |            |     |        |    |      | 59   |
| QY | 1118    |      | CCAAATGATTTGGTACGCAATCTTAAGGTGAACCAATATTTCCATAGAAAGGTGTGAC    |    |            |     |        |    |      | 1177 |
| Db | 60      |      | CCAAATGA--TTGGTACGCAATCTTAACGTGCAACCAATATTTCCAAAGACATGTGTGAC  |    |            |     |        |    |      | 118  |
| QY | 1178    |      | CCATGGCTGTATAAACAAGACTTGGCCCATGTCGCAATGTGACATACCAAGCTTTG      |    |            |     |        |    |      | 1237 |
| Db | 119     |      | CCATGGCTGTTAAGACACAGGACTTGGCCCATGTGCAATGTGACATACCTCAAGC--TTG  |    |            |     |        |    |      | 177  |
| QY | 1238    |      | GAATTGAGGTGATGTTGAAGATGATCAGTGTCTTTTACAAGTCCCTGATCAATGAA      |    |            |     |        |    |      | 1297 |
| Db | 178     |      | GGAAATGAGGTGATGTTGAAGATGATCAGTGTCTTTTACAAGTCCCTGATCAATGAA     |    |            |     |        |    |      | 237  |
| QY | 1298    |      | AATATCTAATATGTGCTCTCTCCCATGGAAGAGATATATCCACAGAGACCGATCATCTGGA |    |            |     |        |    |      | 1357 |
| Db | 238     |      | AATATCTAATATGTGCTCTCTCCCATGGAAGAGATATATCCAGAGACCGATCATCTGGA   |    |            |     |        |    |      | 297  |
| QY | 1358    |      | TATGCTTCAGTACAGGAAACAGATGAACCGCTCTGAGAGAAACGTGCAGTCAACAAT     |    |            |     |        |    |      | 1411 |
| Db | 298     |      | TATGCTTCAGTACAGGAAACAGATGAACCGCTCTGAGAGAAACGTGCAGTCAACAAT     |    |            |     |        |    |      | 357  |

|            |  |        |      |        |                 |
|------------|--|--------|------|--------|-----------------|
| LOCUS      | BG165370   | 966 bp | mRNA | linear | EST 06-FEB-2001 |
| DEFINITION | 602346005F1 NIH_MGC_89 Homo sapiens cDNA IMAGE:4456033 5', |        |      |        |                 |

|           |             |
|-----------|-------------|
| ACCESSION | BG165370    |
| VERSION   | BG165370.1  |
|           | GI:12672073 |

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Cho

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)

Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
Tissue Procurement: ATCC

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found at <http://www.ncbi.nlm.nih.gov/clone/>

Plate: L1AM10250 row: j column: 02  
High quality sequence stop: 649.

```
/organism="Homo sapiens"  
/mol_type="mRNA"
```

```

/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"

```

Site 2: SalI; Cloned unidirectionally; oligo-dT<sup>+</sup> primed. Average insert size 1.3 kb. Library enriched for

|            |       |       |       |       |
|------------|-------|-------|-------|-------|
| BASE COUNT | 275 a | 204 c | 183 g | 244 t |
| ORIGIN,    |       |       |       |       |

|                        |        |                     |        |            |
|------------------------|--------|---------------------|--------|------------|
| Query Match            | 35.1%; | Score 622.4;        | DB 10; | Length 906 |
| Best Local, Similarity | 96.0%; | Pred. No. 7.6e-119; |        |            |





P1 Ford GS, Bloom D, Fairman CG;  
 XX WPI; 2002-055597/07.  
 DR P-PsDB; AAB47800.  
 XX Novel nucleic acid sequences of anergy associated genes, including  
 PT GRAL1 gene useful in the evaluation of pathophysiology or immunotherapy  
 PT of cancer, autoimmune disease and transplant rejection -  
 XX  
 PS Claim 2; Page 45-47; 50pp; English.

XX This sequence encodes human GRAL1 protein. The GRAL1 gene is an  
 CC anergy associated gene which is upregulated during the early stages of  
 CC induction of anergy. GRAL1 has been shown to attenuate IL-2  
 CC transcription in T-cells during response to antigenic stimulation.  
 CC GRAL1 DNA is useful for decreasing the responsiveness of a T cell  
 CC population, especially for synthesis of interleukin-2 (IL-2) in response to  
 CC antigenic stimulation, by up-regulating GRAL1 activity in the T cell  
 CC population. GRAL1 DNA is useful in producing compositions that modulate  
 CC function or maintenance of anergy, for gene therapy, mapping  
 CC functional regions of the encoded protein, to analyse a patient sample  
 CC for the presence of polymorphisms or alterations in expression of  
 CC sequences associated with T cell anergy, disease states, genetic  
 CC predisposition to a disease state, and in studying associated  
 CC physiological pathways. Modulation of the gene activity in vivo is  
 CC useful for prophylactic and therapeutic purposes, such as treating  
 CC autoimmune disease and to enhance immune response to tumour cells, and  
 CC identification of anergic T cells. GRAL1 DNA is useful in the  
 CC evaluation of the pathophysiology or immunotherapy of cancer,  
 CC autoimmune disease, and transplant rejection. Genetic sequences  
 CC involved in anergy induction are useful as markers in the evaluation of  
 CC specific immunotherapies. Functional characterization of genes involved  
 CC in anergy induction allows the elucidation of the mechanisms of T cell  
 CC anergy, including the transcriptional blockade of IL-2, which may be  
 CC manipulated to regulate T cell responses in human disease. Modulation  
 CC of expression of GRAL1 gene is useful in manipulating the anergic  
 CC state. The genetic sequences find use alone or in combinations in  
 CC determining the expression profile of cells relating to anergy, for  
 CC example in screening of candidate biologically active compounds for  
 CC modulation of T cell anergy. Expression of GRAL1 has been found to  
 CC block tyrosine phosphorylation of a protein present in T cells during  
 CC activation.

XX Sequence 1774 BP; 454 A; 432 C; 464 G; 423 T; 1 other;

Query Match 99.9%; Score 1773; DB 24; Length 1774;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACCGCGGTGGCGGCTCTTGAACCTGATGCCCGGCTGCAAG 60  
 Db 1 AGCTGAGCTCCACCGCGGTGGCGGCTCTTGAACCTGATGCCCGGCTGCAAG 60  
 QY 61 AATTCGACAGACCCAGAGAGCTGATCTGGGCAACTGTGCTGACGTCAGTCCCT 120  
 Db 61 AATTCGACAGACCCAGAGAGCTGATCTGGGCAACTGTGCTGACGTCAGTCCCT 120  
 QY 121 CCGGCTCCGACGTAGCTGAGCTCCCAAGTCTCACTCCATTCTTCCCACTGAGCG 180  
 Db 121 CCGGCTCCGACGTAGCTGAGCTCCCAAGTCTCACTCCATTCTTCCCACTGAGCG 180  
 QY 181 GCACCTGCTCAAGACAGAGGCTCTGCAAGGCTAGAGAGGCGCTGCAAGGCGCTGAG 240  
 Db 181 GCACCTGCTCAAGACAGAGGCTCTGCAAGGCTAGAGAGGCGCTGCAAGGCGCTGAG 240  
 QY 241 GGAAGCTGGAGGCGCGCCATGGGCGCGGCTGGGGGCGGGAGTCTCTGCGCGG 300  
 Db 241 GGAAGCTGGAGGCGCGCCATGGGCGCGGCTGGGGGCGGGAGTCTCTGCGCGG 300  
 QY 301 TGGCTGGGCTTTTCAAGATGCTTGGCATGCTGCTTCTGCTGCGCTGAGTCCGAGGCG 360  
 Db 301 TGGCTGGGCTTTTCAAGATGCTTGGCATGCTGCTTCTGCTGCGCTGAGTCCGAGGCG 360

QY 361 ACCCGGTTCCCGGGGGGCTGAAGACAGTGTGACCGCGTACTCAACGTCTCTGGCGGT 420  
 Db 361 ACCCGGTTCCCGGGGGGCTGAAGACAGTGTGACCGCGTACTCAACGTCTCTGGCGGT 420  
 QY 421 TCCGACACGGGAGTGAACCTGATCCGCTGGGACTGAGGAGAGGGGTGTACGGTCA 480  
 Db 421 TCCGACACGGGAGTGAACCTGATCCGCTGGGACTGAGGAGAGGGGTGTACGGTCA 480  
 QY 481 GGACTGCGCGCTGGAGCCTTGGGCTGAGGAGTCTGAGTACCCCGGACGAGCGGCGCT 540  
 Db 481 GGACTGCGCGCTGGAGCCTTGGGCTGAGGAGTCTGAGTACCCCGGACGAGCGGCGCT 540  
 QY 541 TAACGCTGTAAACCCGACACGAAATTCACGAGTCCGACGAGTGGGAGAACCGTGA 600  
 Db 541 TAACGCTGTAAACCCGACACGAAATTCACGAGTCCGACGAGTGGGAGAACCGTGA 600  
 QY 601 AGCTCTTGGTGGGCGCTTCCACACGGGCGGGGGCTGCACTTGGCAGACAAATGCA 660  
 Db 601 AGCTCTTGGTGGGCGCTTCCACACGGGCGGGGGCTGCACTTGGCAGACAAATGCA 660  
 QY 661 TCTGCTTATGAGAGATGGGCGCTGAGACCGTCACTTTAACTTCCCGGACCCGCA 720  
 Db 661 TCTGCTTATGAGAGATGGGCGCTGAGACCGTCACTTTAACTTCCCGGACCCGCA 720  
 QY 721 TGAAGTCAATCCCAATGCTCAACCGGCTGCACTGAGATTTGGATGATGATGGGCA 780  
 Db 721 TGAAGTCAATCCCAATGCTCAACCGGCTGCACTGAGATTTGGATGATGATGGGCA 780  
 QY 781 TCGAAGAGGCAAAAATTCGCAATCTATCAAGAGGCTCAAGTCAATGATGAT 840  
 Db 781 TCGAAGAGGCAAAAATTCGCAATCTATCAAGAGGCTCAAGTCAATGATGAT 840  
 QY 841 AGAAGTGAAGAAAAAATCATGCGCTTGGGAGTCACTATTCATTTTTCGTTCTGT 900  
 Db 841 AGAAGTGAAGAAAAAATCATGCGCTTGGGAGTCACTATTCATTTTTCGTTCTGT 900  
 QY 901 GTCTCTTTTATTTATACGGCGGCAACTGTGGCTATTTATCTTATCTGCTCGAG 960  
 Db 901 GTCTCTTTTATTTATACGGCGGCAACTGTGGCTATTTATCTTATCTGCTCGAG 960  
 QY 961 GCTACGSAATGCAAGACTCAAGAGCAGAGAGGCAATTAAGCAGATGCTAATAA 1020  
 Db 961 GCTACGSAATGCAAGACTCAAGAGCAGAGAGGCAATTAAGCAGATGCTAATAA 1020  
 QY 1021 AGCTATTGGAAGGCTTCACTACGACACTGAAACAGAGACAGAAATTTGGCCCTGA 1080  
 Db 1021 AGCTATTGGAAGGCTTCACTACGACACTGAAACAGAGACAGAAATTTGGCCCTGA 1080  
 QY 1081 TGGAGATGTTGTGCTGTGTGCTGATGATTTGATTAACCAATGATTTGATGCACTCT 1140  
 Db 1081 TGGAGATGTTGTGCTGTGTGCTGATGATTTGATTAACCAATGATTTGATGCACTCT 1140  
 QY 1141 AAGTGCACAACATATTTTCCATTAAGACATGTTTGAACCAATGCTTTAAACACAGAC 1200  
 Db 1141 AAGTGCACAACATATTTTCCATTAAGACATGTTTGAACCAATGCTTTAAACACAGAC 1200  
 QY 1201 TTTGCCCATGTGCAAAATGTGACATCTCAAACTTTGGAAATTTGAGTGAATTTGAAGA 1260  
 Db 1201 TTTGCCCATGTGCAAAATGTGACATCTCAAACTTTGGAAATTTGAGTGAATTTGAAGA 1260  
 QY 1261 TGGATCAAGTCTTTAACAAGTCCCTGTATCCATGAATAATCTAATAGTCCCTCCCA 1320  
 Db 1261 TGGATCAAGTCTTTAACAAGTCCCTGTATCCATGAATAATCTAATAGTCCCTCCCA 1320  
 QY 1321 TGAAGAGATTAATGCAAGCGACGACCCGATCATCTGATATGCTTCACTACAGAGAACGA 1380  
 Db 1321 TGAAGAGATTAATGCAAGCGACGACCCGATCATCTGATATGCTTCACTACAGAGAACGA 1380  
 QY 1381 TGAACCGCTCTGAGAGAACACCTGAGCTCAACATGAAATGATCTACAGCTGTAAACA 1440  
 Db 1381 TGAACCGCTCTGAGAGAACACCTGAGCTCAACATGAAATGATCTACAGCTGTAAACA 1440  
 QY 1441 TGAAGCAATTCGTGGAGTGTGATGTTATTCCTCATGATGACACCAACCTTTGAAGA 1500

```

Db      1441 TGAAGCAATTCGTGGGAGGTGATGTATTCTCATGTGNGACACCAACCTTTGAGA 1500
Qy      1501 AGACGAACTCTCTAATCAAGAGCTGCTGTTCGAGAAATTAATCTTAATCTGTGTA 1560
Db      1501 AGACGAACTCTCTAATCAAGAGAGCTGCTGTTCGAGAAATTAATCTTAATCTGTGTA 1560
Qy      1561 ATAGAAACCTGACCACTTATGAAATTAACAGACTGCGCAATCGAGGCTCAATTTCTATTAA 1620
Db      1561 ATAGAAACCTGACCACTTATGAAATTAACAGACTGCGCAATCGAGGCTCAATTTCTATTAA 1620
Qy      1621 TAAATGATTAATTTAATTAATAATTAAGAGTATCTGAAAGTCTCGATGATATAT 1680
Db      1621 TAAATGATTAATTTAATTAATAATTAAGAGTATCTGAAAGTCTCGATGATATAT 1680
Qy      1681 ATGCTATAGCTTAAAGCTTAAATTAATTTAACTGCTTAACTTTTTCACCAACCTCA 1740
Db      1681 ATGCTATAGCTTAAAGCTTAAATTAATTTAACTGCTTAACTTTTTCACCAACCTCA 1740
Qy      1741 TTATATATATTTTTCATAGGCAAGTTCTCTCAG 1774
Db      1741 TTATATATATTTTTCATAGGCAAGTTCTCTCAG 1774

```

## RESULT 2

AA15995  
ID AA15995 standard; cDNA; 2773 BP.

XX AA15995;

DT 12-JUN-2000 (first entry)

DE Human protein clone HP10574 full length coding sequence.

XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;  
 KW cytokine production; cell proliferation; cell differentiation;  
 KW immune deficiency; infectious disease; autoimmune disorder; asthma;  
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;  
 KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;  
 KW nervous system disorder; Alzheimer's disease; Parkinson's disease;  
 KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;  
 KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;  
 KW coagulation disorder; myocardial infarction; inflammatory condition;  
 KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;  
 KW nephritis; therapy; ss.

XX Homo sapiens.

XX PN W0200005367-A2.

XX 03-FEB-2000.

PF 22-JUL-1999; 99WO-JP03929.

XX 24-JUL-1998; 98JP-0208820.

PR 07-AUG-1998; 98JP-0224105.

PR 25-AUG-1998; 98JP-0238116.

PR 09-SEP-1998; 98JP-0254736.

PR 23-SEP-1998; 98JP-0275505.

XX (SAGA) SAGAMI CHEM RES CENT.

XX (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;

XX MPI; 2000-182694/16.

XX P-PSDB; AAY94897.

XX Novel human proteins having hydrophobic domains useful for treating

XX osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,

XX multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -

XX This sequence encodes a human protein of the invention, which has hydrophobic domains. The DNA sequences can be used as a probe or as a genetic marker. The protein can also be used as a marker and to identify potential genetic disorders. The DNA and protein can also be used as nutritional sources or supplements. The protein exhibits cytokine, cell proliferation, cell differentiation activities and induces production of other cytokines in certain cell populations. The protein also exhibits immune stimulating or immune suppressing activity. It can be used in the treatment of various immune deficiencies and disorders, and to treat infectious diseases caused by viral, bacterial, fungal or other infections. The protein is also used for treating autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid arthritis. It is also useful in the treatment of allergic reactions and conditions such as asthma, and in immune suppression after organ transplantation. The protein is useful in regulation of haematopoiesis and consequently in the treatment of myeloid or lymphoid cell deficiencies. It is also used in compositions for tissue growth or regeneration. The protein is also used in the treatment of osteoporosis or osteoarthritis and in the treatment of periodontal disease and other tooth repair processes. The protein is used in the treatment of nervous system disorders such as Alzheimer's disease, Parkinson's disease, and Huntington's disease. They are useful for protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage. They are also used for promoting or inhibiting tissue differentiation. They are also used as contraceptives since they exhibit activin or inhibin related activities and as a fertility inducing therapeutic. They are used for treating various coagulation disorders and in treatment and prevention of conditions resulting from coagulation activities e.g. myocardial infarction or stroke. They also acts as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. They are used to treat inflammatory conditions such as septic shock, sepsis, ischaemia reperfusion injury, arthritis, and nephritis. They can be used to prevent tumours.

Sequence 2773 BP; 803 A; 545 C; 617 G; 808 T; 0 other;

Query Match 93.7%; Score 1662.8; DB 21; Length 2773;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```

Qy      74 CCGAGAGCTGCACTGCGGCAACCTGTGTCTGACGCTACGCTCTGCTGCTCCAG 133
Db      22 CCGAGAGCTGCACTGCGGCAACCTGTGTCTGACGCTACGCTCTGCTGCTCCAG 81
Qy      134 TAGCTGCACTGCTCCAGTCTCACTCACTTCCTTCCCACTTGGCGGCACTGCTCAAG 193
Db      82 TAGCTGCACTGCTCCAGTCTCACTCACTTCCTTCCCACTTGGCGGCACTGCTCAAG 141
Qy      194 ACCAGAGTCCGCAAGCGCTGAGAGGCGGTGCGAGGAGCGCTAGGAGACCTGCGAGC 253
Db      142 ACCAGAGTCCGCAAGCGCTGAGAGGCGGTGCGAGGAGCGCTAGGAGACCTGCGAGC 201
Qy      254 GCGGCGCATGAGGAGCGCGCGCTGAGGAGCGGAGTCTCTGCGCGGTGCTGCGGCTTT 313
Db      202 GCGGCGCATGAGGAGCGCGCGCTGAGGAGCGGAGTCTCTGCGCGGTGCTGCGGCTTT 261
Qy      314 TCCAGATTGCGGAGATGAGTCTTCTGCTGAGCGCTGAGTCCGAGGACCCGAGTCCCGG 373
Db      262 TCCAGATTGCGGAGATGAGTCTTCTGCTGAGCGCTGAGTCCGAGGACCCGAGTCCCGG 321
Qy      374 GGGGCTGAAGAGTGTGACCCGAGTACCTCAAGTGTCTTGTGAGGAGTCCGAGCAAGGGA 433
Db      322 GGGGCTGAAGAGTGTGACCCGAGTACCTCAAGTGTCTTGTGAGGAGTCCGAGCAAGGGA 381
Qy      434 GTGACCTGTAAGTGTGTGAGAGCTGAGCGAGAGAGGCGTGTACCGGACGATCTGCGCTG 493
Db      382 GTGACCTGTAAGTGTGTGAGAGCTGAGCGAGAGAGGCGTGTGTACCGGACGATCTGCGCTG 441
Qy      494 GAGCCTGTGAGTGTGAGTCTGTGTAACGCGCGAGCGAGCGGCGCTTACGCGCTGTAC 553
Db      442 GAGCCTGTGAGTGTGAGTCTGTGTAACGCGCGAGCGAGCGGCGCTTACGCGCTGTAC 501

```



QY 554 CCGGACACGATTTTACAGGTCGCCAGGTTTGGGGAAGCAACGTCAGTCTCTGTTG 613  
 DB 502 CCGGACACGATTTTACAGGTCGCCAGGTTTGGGGAAGCAACGTCAGTCTCTGTTG 561  
 QY 614 GCCCTCATCCAAACGAGCGGCGGCGCTGCACTTGGCAGACAAAGATCATCTGGCTTATG 673  
 DB 562 GCCCTCATCCAAACGAGCGGCGGCGCTGCACTTGGCAGACAAAGATCATCTGGCTTATG 621  
 QY 674 AGATGGGCGCTTGGAGCCGTCATCTTTAATCTTCCCGGGAACCCGATAGGTCATCCCC 733  
 DB 622 AGAGGGCGCTTGGAGCCGTCATCTTTAATCTTCCCGGGAACCCGATAGGTCATCCCC 681  
 QY 734 ATGTCTACACCGGCTGAGTGAACATTTGTGCAATCATGTGGCAATCGAAGGACA 793  
 DB 682 ATGTCTACACCGGCTGAGTGAACATTTGTGCAATCATGTGGCAATCGAAGGACA 741  
 QY 794 AAAATTCGCAATCTAATTTCAAGAGGCAACAAGTGAACAAAGTCATAGAGGAAA 853  
 DB 742 AAAATTCGCAATCTAATTTCAAGAGGCAACAAGTGAACAAAGTCATAGAGGAAA 801  
 QY 854 AAAAGTGGCCCTGGGATGATCAATTTCTTTTGTCTGTCGTCCTTTTAT 913  
 DB 802 AAACATGGCCCTGGGATGATCAATTTCTTTTGTCTGTCGTCCTTTTAT 861  
 QY 914 ATTAAGCGGCAACGTGGCTATTTTATCTTTTATCTGTCGAAAGCTACGGAATGCA 973  
 DB 862 ATTAAGCGGCAACGTGGCTATTTTATCTTTTATCTGTCGAAAGCTACGGAATGCA 921  
 QY 974 AGAGCTCAAGCGAGAGAGAGAGGCAATTAAGCGAGTCATTAAGGCAATTTGAAG 1033  
 DB 922 AGAGCTCAAGCGAGAGAGAGAGGCAATTAAGCGAGTCATTAAGGCAATTTGAAG 981  
 QY 1034 CTTCACACTGACCACTGTAACAAAGAGACAGAGAAATGGCCCTGATGAGATAGTGT 1093  
 DB 982 CTTCACACTGACCACTGTAACAAAGAGACAGAGAAATGGCCCTGATGAGATAGTGT 1041  
 QY 1094 GCTGTGTCATTTGATTTGATTAACCAATGATTTGGTACGATCTTAAGTGAACCT 1153  
 DB 1042 GCTGTGTCATTTGATTTGATTAACCAATGATTTGGTACGATCTTAAGTGAACCT 1101  
 QY 1154 ATTTCCATAAGCATGTGTTGACCCATGCTGTTAAACACAAAGACTTCCCATATGTC 1213  
 DB 1102 ATTTCCATAAGCATGTGTTGACCCATGCTGTTAAACACAAAGACTTCCCATATGTC 1161  
 QY 1214 AAATGTGACATCTCAAGCTTTGGGAATGAGGTGATGTTGAAGATGATGATGTC 1273  
 DB 1162 AAATGTGACATCTCAAGCTTTGGGAATGAGGTGATGTTGAAGATGATGATGTC 1221  
 QY 1274 TTACAGATCCCTGATGCAATGAAATATCTAATATGAGTCCCTCCCATGAAGAGATAT 1333  
 DB 1222 TTACAGATCCCTGATGCAATGAAATATCTAATATGAGTCCCTCCCATGAAGAGATAT 1281  
 QY 1334 CCGACGAGACCCGATCATCTGATATGCTTCAATGACAGGAAACAGATGAACCCCTCTG 1393  
 DB 1282 CCGACGAGACCCGATCATCTGATATGCTTCAATGACAGGAAACAGATGAACCCCTCTG 1341  
 QY 1394 GAGGACACCTGACATCAACAATGAAGTCAAGCTGTTAAACCAAGAGCAAAATCTC 1453  
 DB 1342 GAGGACACCTGACATCAACAATGAAGTCAAGCTGTTAAACCAAGAGCAAAATCTC 1401  
 QY 1454 GTGGCAGTGAATTTATCTCATGTGACCAACCAACCTTTGAAGAAAGCAAACTCTC 1513  
 DB 1402 GTGGCAGTGAATTTATCTCATGTGACCAACCAACCTTTGAAGAAAGCAAACTCTC 1461  
 QY 1514 AATCAAGAGACTGTGTGAGAAATTAATCTTAATGTGTAAATTAAGAACTTGA 1573  
 DB 1462 AATCAAGAGACTGTGTGAGAAATTAATCTTAATGTGTAAATTAAGAACTTGA 1521  
 QY 1574 ACCATTAATTAACAGAACTGCAATCAGGCTTATTTCTAATTAATTAATTAATTA 1633  
 DB 1522 ACCATTAATTAACAGAACTGCAATCAGGCTTATTTCTAATTAATTAATTAATTA 1581

QY 1634 TTATAATAATTAAGAGTGAATCTGAAGAGTCTGAGATGATTAATTAATTAATTA 1693  
 DB 1582 TTATAATAATTAAGAGTGAATCTGAAGAGTCTGAGATGATTAATTAATTAATTA 1641  
 QY 1694 AATGGCTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 1753  
 DB 1642 A--TGGCTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 1698  
 QY 1754 CATAGGCAAGTTCTCTCAG 1774  
 DB 1699 CATAGGCAAGTTCTCTCAG 1719

RESULT 3  
 AAI72321  
 ID AAI72321 standard; cDNA; 2796 BP.

AC AAI72321;  
 DT 15-APR-2002 (first entry)  
 DE ISIGP-4 cDNA.  
 XX Human; intracellular signalling protein; ISIGP; gene; cell proliferation;  
 XX autoimmunity; inflammation; gastrointestinal disorder;  
 KW reproductive disorder; developmental disorder; ss.  
 OS Homo sapiens.

Key location/Qualifiers  
 FT CDS 217..1503  
 FT /\*tag= a  
 FT /product= "ISIGP-4"

W0200194391-A2.  
 PD 13-DEC-2001.

PF 07-JUN-2001; 2001WO-US18595.

PR 08-JUN-2000; 2000US-210582P.  
 PR 16-JUN-2000; 2000US-212443P.

PA (INCY-) INCYTE GENOMICS INC.

PI Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT;  
 PI Xu Y, Hafalia A, Azimzai Y, Walia NK;

DR WPI; 2002-154564/20.  
 DR P-PSDB; AAB47874.

PT New human intracellular signalling protein and polynucleotides useful  
 PT for diagnosing, treating or preventing cell proliferative;  
 PT autoimmune/inflammatory, gastrointestinal, reproductive and  
 PT developmental disorders -  
 PS Claim 5; Page 104; 106pp; English.

CC The sequences given in AAI72318-22 encode novel human intracellular  
 CC signalling proteins (ISIGP). The polynucleotides and ISIGP proteins  
 CC may be used for the diagnosis, treatment or prevention of cell  
 CC proliferative, autoimmune/inflammatory, gastrointestinal, reproductive  
 CC and developmental disorders. The protein encoded by this sequence has  
 CC homology to mouse gl-related zinc finger protein.

SQ Sequence 2796 BP; 821 A; 548 C; 620 G; 807 T; 0 other;

Query Match 93.7%; Score 1662.8; DB 24; Length 2796;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 74 CCGAGAGGTCATCTCGGCAACCTGATGTCAGCTTACCTCTGAGCTCCGACG 133





CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid

Db 601 CTTGGGTGATCACTATTCATTTTTCGTTTCGTGTCCTTTTATTAATTACGGCG 660

```
QY 923 GCAACTGAGGCTAATTTATCTTTATTCGCTGAGGCTACGAAAGCAGCTCA 982
DB 661 GCAACTGAGGCTAATTTATCTTTATTCGCTGAGGCTACGAAAGCAGCTCA 720
QY 983 AGCAGAGAGCAGAGCAGATTAAGCAGAGCTTAAAGATTTGAGAGCTTCACTA 1042
DB 721 AGCAGAGAGCAGAGCAGATTAAGCAGAGCTTAAAGATTTGAGAGCTTCACTA 780
QY 1043 CGCAGACTGAAACAGAGAGCAGAGATTTGGCCCTGATGAGATGTTGCTGTGTC 1102
DB 781 CGCAGACTGAAACAGAGAGCAGAGATTTGGCCCTGATGAGATGTTGCTGTGTC 840
QY 1103 ATTGAATGTATTAACCAATATTTGTTGACGATCTTTAGCTGACATTTTCCAT 1162
DB 841 ATTGAATGTATTAACCAATATTTGTTGACGATCTTTAGCTGACATTTTCCAT 900
QY 1163 AAGCAGATGTTGACCCATGCTGTTAAACAGACAGCTTGGCCCATGTCATATGTGC 1222
DB 901 AAGCAGATGTTGACCCATGCTGTTAAACAGACAGCTTGGCCCATGTCATATGTGC 960
QY 1223 ATACTCAAGCTTTGGGATTTAGGTGATGTTGTAAGATGATCACTGCTTTCAAGTC 1282
DB 961 ATACTCAAGCTTTGGGATTTAGGTGATGTTGTAAGATGATCACTGCTTTCAAGTC 1020
QY 1283 CCTGATCCATGAAATATCTATATAGTCCCTCCCTCCCATGAGAGATTAATCCAGCGAG 1342
DB 1021 CCTGATCCATGAAATATCTATATAGTCCCTCCCTCCCATGAGAGATTAATCCAGCGAG 1080
QY 1343 ACCGATATCTTGGATTTGCTTCACTACAGAGAAACAGATGAAACGCTCTGAGAGAACAC 1402
DB 1081 ACCGATATCTTGGATTTGCTTCACTACAGAGAAACAGATGAAACGCTCTGAGAGAACAC 1140
QY 1403 GTGAGTCAACCAATGAAAGTCTACAGCTGTAAACCATGACGAAATCTGTGAGCAGTG 1462
DB 1141 GTGAGTCAACCAATGAAAGTCTACAGCTGTAAACCATGACGAAATCTGTGAGCAGTG 1200
QY 1463 GATGTATTTCTCTATGAGACCAACCACTTGAAGAGAGAAATCTCTATCAAGAG 1522
DB 1201 GATGTATTTCTCTATGAGACCAACCACTTGAAGAGAGAAATCTCTATCAAGAG 1260
QY 1523 ACTGCTGTTCCAGAAATTAATCT 1546
DB 1261 ACTGCTGTTCCAGAAATTAATCT 1284

RESULT 6
AAS25884
ID AAS25884 standard; cDNA; 2306 BP.
XX
AC AAS25884;
XX
DT 07-NOV-2001 (first entry)
XX
DB Human cDNA encoding a novel secreted protein, Seq ID 63.
XX
XX Human; immunosuppressive; antiarthritic; ss; antineoplastic;
XX cyostatic; cardiant; vasotropic; cerebroprotective; nootropic;
XX neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
XX vulnary; secreted protein; rheumatoid arthritis;
XX hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
XX cerebrovascular disorder; cerebral ischemia; angiogenesis;
XX nervous system disorder; Alzheimer's disease; infection; ocular disorder;
XX corneal infection; wound healing; epithelial cell proliferation;
XX skin aging; food additive; preservative; antiproliferative.
OS Homo sapiens.
XX
XX W0000155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001W0-US01341.
XX
```

```
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
```

PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 05-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM,  
 XX  
 XX WPI; 2001-488783/53.  
 XX P-PSDB; AAU15897.  
 XX  
 XX New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and

PT used as food additives or preservatives -  
 XX  
 XX Claim 1; SEQ ID No 63; 980bp; English.  
 XX

CC The invention relates to isolated nucleic acid molecules and proteins are used to  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. Antibodies to the proteins can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angioneurosis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection, and many other  
 CC disorders listed in the specification. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence encodes a novel secreted protein of the invention.

Query Match 66.4%; Score 1178.8; DB 22; Length 2306;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

|    |      |   |      |
|----|------|---|------|
| QY | 547  | CTGTAACCCGCAACAGATTTCACGGTGGCCGTTGGGGAAGACCCGTCAGATCTC    | 606  |
| Db | 12   | CTGTAACCCGCAACAGATTTCACGGTGGCCGTTGGGGAAGACCCGTCAGATCTC    | 71   |
| QY | 607  | TTGGTTGGCCCTCAATCCAAAGCGCGGGGCTGCACTTGGCGAATAATCCATCTGCG  | 666  |
| Db | 72   | TTGGTTGGCCCTCAATCCAAAGCGCGGGGCTGCACTTGGCGAATAATCCATCTGCG  | 131  |
| QY | 657  | TTATGAGAGATGGGGCTGTGAGACCGTCACTTTAACTCCCGGGAACCCGGAATGAGT | 726  |
| Db | 132  | TTATGAGAGAGGGGGCTGTGAGACCGTCACTTTAACTCCCGGGAACCCGGAATGAGT | 191  |
| QY | 727  | CATCCCAATGCTCAACCGGGGTGATGAGATTGGCAATCATGATGGCAATCTGAA    | 786  |
| Db | 132  | CATCCCAATGCTCAACCGGGGTGATGAGATTGGCAATCATGATGGCAATCTGAA    | 251  |
| QY | 787  | AGGCAAAAATTTGCAATCTTTCAAAGAGGCAATCAAGTCAATGCTAAGAGT       | 846  |
| Db | 252  | AGGCAAAAATTTGCAATCTTTCAAAGAGGCAATCAAGTCAATGCTAAGAGT       | 311  |
| QY | 847  | AGGCAAAAATTTGCAATCTTTCAAAGAGGCAATCAAGTCAATGCTAAGAGT       | 905  |
| Db | 312  | AGGCAAAAATTTGCAATCTTTCAAAGAGGCAATCAAGTCAATGCTAAGAGT       | 371  |
| QY | 906  | TTTTTATTTATGCGCGGCACTGTGGGCTATTTTATTTTATTTTGTGCAAGGCTAC   | 965  |
| Db | 372  | TTTTTATTTATGCGCGGCACTGTGGGCTATTTTATTTTATTTTGTGCAAGGCTAC   | 431  |
| QY | 966  | GGAATGCAAGAGCTCAAGAGGCAATCAAGGCAATTTAAAGGAGATGCTAAGAGT    | 1025 |
| Db | 432  | GGAATGCAAGAGCTCAAGAGGCAATCAAGGCAATTTAAAGGAGATGCTAAGAGT    | 491  |
| QY | 1026 | TTGGAAGGCTCAATCAAGGCAATCAAGGCAATTTAAAGGAGATGCTAAGAGT      | 1085 |
| Db | 492  | TTGGAAGGCTCAATCAAGGCAATCAAGGCAATTTAAAGGAGATGCTAAGAGT      | 551  |
| QY | 1086 | ATAAGTGTCTGTGATGATTAATTTATTAACCAATTTTGGACGATCTTAAGT       | 1145 |
| Db | 552  | ATAAGTGTCTGTGATGATTAATTTATTAACCAATTTTGGACGATCTTAAGT       | 611  |

QY 1146 GCAACCAATATTTTCATAGACATGTGTGACCCATGGCTGTTAAACACAGACCTGCC 1205  
 |||||  
 DB 612 GCAACCAATATTTTCATAGACATGTGTGACCCATGGCTGTTAAACACAGACCTGCC 671  
 |||||  
 QY 1206 CCAATGTGCAATGTGCATCTCAACAGCTTTGGGAATGAGGTGAGATGAT 1265  
 |||||  
 DB 672 CCAATGTGCAATGTGCATCTCAACAGCTTTGGGAATGAGGTGAGATGAT 731  
 |||||  
 QY 1266 CAGATGCTTTTACAGTCCCTGATCCATGAATATATATAGTGCCTCCCTCCATGAG 1325  
 |||||  
 DB 732 CAGATGCTTTTACAGTCCCTGATCCATGAATATATATAGTGCCTCCCTCCATGAG 791  
 |||||  
 QY 1326 AGGATATCCAGCCGAGACCCGATCTCTGGAATGCTTCAGTACAGGAAACGATGAC 1385  
 |||||  
 DB 792 AGGATATCCAGCCGAGACCCGATCTCTGGAATGCTTCAGTACAGGAAACGATGAC 851  
 |||||  
 QY 1386 CGCCTCTGAGAGAACGTCGACGTCAACAAATGAAGTCTACAGCTGTAAACGATGAG 1445  
 |||||  
 DB 852 CGCCTCTGAGAGAACGTCGACGTCAACAAATGAAGTCTACAGCTGTAAACGATGAG 911  
 |||||  
 QY 1446 CAAATCTGTGCGAGGAGATGTATCTCTCATGTGACACACCCATCTTTGAGAGAGCG 1505  
 |||||  
 DB 912 CAAATCTGTGCGAGGAGATGTATCTCTCATGTGACACACCCATCTTTGAGAGAGCG 971  
 |||||  
 QY 1506 AAACCTCAATCAAGAGACTGCTGTGAGAAATTAATCTTAAATCTGTGAATGAG 1565  
 |||||  
 DB 972 AAACCTCAATCAAGAGACTGCTGTGAGAAATTAATCTTAAATCTGTGAATGAG 1031  
 |||||  
 QY 1566 AAACCTGAACCATTAATTAACGAACTCCGATCAGAGGCTTCTTCAATTAAT 1625  
 |||||  
 DB 1032 AAACCTGAACCATTAATTAACGAACTCCGATCAGAGGCTTCTTCAATTAAT 1091  
 |||||  
 QY 1626 TCGAATAATTTAAATTAATTAAGTGAATCTGGAAGTGTGAGTAAATATATAGCT 1685  
 |||||  
 DB 1092 TCGAATAATTTAAATTAATTAAGTGAATCTGGAAGTGTGAGTAAATATAGCT 1151  
 |||||  
 QY 1686 ATAGTTAAATATGCTTAAATTAATTAACCTGTAATCTTTTCCACCAACCTATTA 1745  
 |||||  
 DB 1152 ATAGTTAAATATGCTTAAATTAATTAACCTGTAATCTTTTCCACCAACCTATTA 1208  
 |||||  
 QY 1746 ATATTTTCAATAGGCAAGTTTCTCTCAG 1774  
 |||||  
 DB 1209 ATATTTTCAATAGGCAAGTTTCTCTCAG 1237  
 |||||  
 RESULT 7  
 ABX73225  
 ID ABX73225 standard; DNA; 2306 BP.  
 AC ABX73225;  
 XX  
 DT 18-MAR-2003 (first entry)  
 XX  
 DE Human novel polynucleotide #53.  
 XX  
 XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
 XX muscular disorder; respiratory disease; reproductive disorder;  
 XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 XX hyperproliferative disorder; inflammatory disease; allergic reaction;  
 XX blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 XX hemostatic; antiarteriosclerotic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US200212753-A1.  
 XX  
 PD 19-SEP-2002.  
 XX  
 XX 17-JAN-2001; 2001US-0764864.  
 XX  
 XX 31-JAN-2000; 2000US-179065P.  
 XX  
 XX 04-FEB-2000; 2000US-180628P.  
 XX

PR 28-JUN-2000; 2000US-214886P.  
 PR 07-JUL-2000; 2000US-216647P.  
 PR 07-JUL-2000; 2000US-216880P.  
 PR 11-JUL-2000; 2000US-217487P.  
 PR 11-JUL-2000; 2000US-217966P.  
 PR 14-JUL-2000; 2000US-218290P.  
 PR 26-JUL-2000; 2000US-220963P.  
 PR 26-JUL-2000; 2000US-220964P.  
 PR 14-AUG-2000; 2000US-224518P.  
 PR 14-AUG-2000; 2000US-224519P.  
 PR 14-AUG-2000; 2000US-22567P.  
 PR 14-AUG-2000; 2000US-22568P.  
 PR 14-AUG-2000; 2000US-22570P.  
 PR 14-AUG-2000; 2000US-225447P.  
 PR 14-AUG-2000; 2000US-225757P.  
 PR 14-AUG-2000; 2000US-22578P.  
 PR 22-AUG-2000; 2000US-226868P.  
 PR 30-AUG-2000; 2000US-228924P.  
 PR 01-SEP-2000; 2000US-229287P.  
 PR 01-SEP-2000; 2000US-229343P.  
 PR 01-SEP-2000; 2000US-229344P.  
 PR 01-SEP-2000; 2000US-229345P.  
 PR 05-SEP-2000; 2000US-229509P.  
 PR 05-SEP-2000; 2000US-229513P.  
 PR 08-SEP-2000; 2000US-231413P.  
 PR 21-SEP-2000; 2000US-234423P.  
 PR 21-SEP-2000; 2000US-234474P.  
 PR 25-SEP-2000; 2000US-234997P.  
 PR 27-SEP-2000; 2000US-235834P.  
 PR 29-SEP-2000; 2000US-236327P.  
 PR 29-SEP-2000; 2000US-236367P.  
 PR 29-SEP-2000; 2000US-236368P.  
 PR 29-SEP-2000; 2000US-236369P.  
 PR 29-SEP-2000; 2000US-236370P.  
 PR 02-OCT-2000; 2000US-236802P.  
 PR 02-OCT-2000; 2000US-237037P.  
 PR 02-OCT-2000; 2000US-237038P.  
 PR 02-OCT-2000; 2000US-237039P.  
 PR 02-OCT-2000; 2000US-237040P.  
 PR 13-OCT-2000; 2000US-239355P.  
 PR 20-OCT-2000; 2000US-240960P.  
 PR 20-OCT-2000; 2000US-241785P.  
 PR 20-OCT-2000; 2000US-241809P.  
 PR 01-NOV-2000; 2000US-244617P.  
 PR 17-NOV-2000; 2000US-249299P.  
 PR 08-DEC-2000; 2000US-251866P.  
 PR 08-DEC-2000; 2000US-251868P.  
 PR 08-DEC-2000; 2000US-251869P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBEN/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PT Rosen CA, Ruben SM, Barash SC;  
 XX  
 XX WFI; 2003-14744/14.  
 DR P-PSDB; ABUS4965.  
 XX  
 PT New polypeptides and nucleic acids, useful in gene therapy for  
 PT treating, inhibiting or preventing e.g. neural, immune system,  
 PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,  
 PT cardiovascular or renal disorders -  
 XX  
 XX Claim 1; SEQ ID NO 63; 402PD; English.  
 PS  
 XX  
 XX The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid  
 CC arthritis and multiple sclerosis), muscular disorders, respiratory  
 CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),  
 CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,  
 CC cardiovascular disorders (e.g. congenital heart defects, Ebstein's



anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute  
kidney failure and end-stage renal disease), hyperproliferative disorders  
(e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g.  
septic shock, bursitis and appendicitis), allergic reactions and  
conditions (e.g. asthma), blood related disorders (e.g. thrombosis,  
thrombocytosis and myocardial infarction) and cancerous disorders.  
Sequences ABX73173-ABX74167 represent human novel polynucleotides of the  
invention.

50 Sequence 2306 BP; 753 A; 391 C; 438 G; 724 T; 0 other;

|                            |        |               |           |              |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match                | 66.4%; | Score 1178.8; | DB 25;    | Length 2306; |
| Best Local Similarity      | 99.0%; | Pred. No. 0;  |           |              |
| Matches 1217; Conservative | 0;     | Mismatches 8; | Indels 4; | Gaps 3;      |

|    |      |   |      |
|----|------|---|------|
| QY | 547  | TTGTAACCCGACACGAATTTTCAAGGTGACCAGGTGTTGGGAAAGACCCGTCGAAGCTC     | 606  |
| Db | 12   | CTGTAAACCGGACACGAATTTCAAGGTGACCAGGTGTTGGGAAAGACCCGTCGAAGCTC     | 71   |
| QY | 607  | TTGGTGGCCCTCACTCCAAAGCGGCGGGGGCTGCACCTTCCGACACAAATCCATCTGCG     | 666  |
| Db | 72   | TTGGTGGCCCTCACTCCAAAGCGGCGGGGGCTGCACCTTCCGACACAAATCCATCTGCG     | 131  |
| QY | 667  | TTATGAGAGATGAGCGCTGTGAGACCGCTCATCTTTAACTTCCCGGACCCGACATGAGGT    | 726  |
| Db | 132  | TTATGAGAGAGAGCGCGCTGTGAGACCGCTCATCTTTAACTTCCCGGACCCGACATGAGGT   | 191  |
| QY | 727  | CATCCCAAGTGTCAACCCGSGTGACGTGACATTTGTCGATCATGATGCGCATCTGAA       | 786  |
| Db | 192  | CATCCCAAGTGTCAACCCGSGTGACGTGACATTTGTCGATCATGATGCGCATCTGAA       | 251  |
| QY | 787  | AGGCACAAAATCTGCAATCTATTCAAAGGCGATACAACTGACAAATGTCATGAGAT        | 846  |
| Db | 252  | AGGCACAAAATCTGCAATCTATTCAAAGGCGATACAACTGACAAATGTCATGAGAT        | 311  |
| QY | 847  | AGGGAAAAAATCATGGCCCTGGGTGAATCACTATTCAA-TTTTTTCGTTTGTGTCCT       | 905  |
| Db | 312  | AGGGAAAAAATCATGGCCCTGGGTGAATCACTATTCAA-TTTTTTCGTTTGTGTCCT       | 371  |
| QY | 906  | TTTTTTATTATTCGCGGCGACACTGTTGGGCTATTTTATCTTTATCTGCTCGAAGCTAC     | 965  |
| Db | 372  | TTTTTTATTATTCGCGGCGACACTGTTGGGCTATTTTATCTTTATCTGCTCGAAGCTAC     | 431  |
| QY | 966  | GGATGCAAGAGCTCAAAGAGAGAAAGCAGAGGCAATTTAAAGCAGATGCTTAAAAAACCTA   | 1026 |
| Db | 432  | GGATGCAAGAGCTCAAAGAGAGAGAAAGCAGAGGCAATTTAAAGCAGATGCTTAAAAAACCTA | 491  |
| QY | 1026 | TTGGAAAGGCTTCATCTACGACACTGAAAAAAGAGAAACAAAGAAATTTGGCCCTGATGGAG  | 1089 |
| Db | 492  | TTGGAAAGGCTTCATCTACGACACTGAAAAAAGAGAAACAAAGAAATTTGGCCCTGATGGAG  | 551  |
| QY | 1086 | ATATGTTGTGCTGTGTCATTTGAATTTGATTAACCAAAATGATTTGTGACGATCTTAAAGT   | 1144 |
| Db | 552  | ATATGTTGTGCTGTGTCATTTGAATTTGATTAACCAAAATGATTTGTGACGATCTTAAAGT   | 611  |
| QY | 1146 | GCAACACATTTTTCCATAAGACACTGTGTGTACCCATGCTGTTTAAACCAAGACTGGCC     | 1205 |
| Db | 612  | GCAACACATTTTTCCATAAGACACTGTGTGTACCCATGCTGTTTAAACCAAGACTGGCC     | 671  |
| QY | 1206 | CGATGTGCAAAATGACATATCTCAAAAGCTTTGGGAATTTGAAGTGATGTTGTAAGATGGAT  | 1265 |
| Db | 672  | CGATGTGCAAAATGACATATCTCAAAAGCTTTGGGAATTTGAAGTGATGTTGTAAGATGGAT  | 731  |
| QY | 1266 | CAGTGTCTTTTCAAGTCCCTGTATCCATGATGAATCTATATATGTCCTCTCCCATGAG      | 1325 |
| Db | 732  | CAGTGTCTTTTCAAGTCCCTGTATCCATGATGAATCTATATATGTCCTCTCCCATGAG      | 791  |
| QY | 1326 | AGGATATAGCAGAGGACACCGGACATCTGGAATATGCTTCAGTACAGAGGAAACAGATGAA   | 1389 |
| Db | 792  | AGGATATAGCAGAGGACACCGGACATCTGGAATATGCTTCAGTACAGAGGAAACAGATGAA   | 851  |
| QY | 1386 | CGGCTCTGGAGAAACATGTGCACTCAACAAATGAAAGTCACAGCTGTGTAACCATGAA      | 1445 |

|          |  |  |      |
|----------|--|--|------|
| Db       | 852  | CGCTCTGTGAGAAACAAGTGCAGTGCACAAAGAAAGTCTACAGCTGGTAAACATGAG  | 911  |
| Qy       | 1446   | CAATTCCTGTGGCAGTGAATGTATTCTCATGTGACAACTCACCTTTGAGAAAGAGC   | 1505 |
| Db       | 912  | CAATTCGTGTGGCAGTGAATGTATTCTCATGTTGCAACCCCAACTTTTAAAGAGAG   | 971  |
| Qy       | 1506   | AAACTCTCAATCAGAGACGTGCTGTTCAGAAATTAATCTTAAATCTGTGAAATAGA   | 1565 |
| Db       | 972  | AAACTCTCAATCAGAGAGACGTGCTGTTCAGAAATTAATCTTAAATCTGTGAAATAGA | 1031 |
| Qy       | 1566   | AAACTTGAACCATTTAGTAATATACGAATCTGCAATCAGAGGCTTGTCTTAAAT     | 1625 |
| Db       | 1032   | AAACTTGAACCATTTAGTAATATACGAATCTGCAATCAGAGGCTTGTCTTAAAT     | 1091 |
| Qy       | 1626   | TGCAATAATTTAAATAAATPAGAGTGAATCTGAAAGTGTCAATGATCTAATATTAGCT | 1685 |
| Db       | 1092   | TGCAATAATTTAAATAAATPAGAGTGAATCTGAAAGTGTCAATGATCTAATATTAGCT | 1151 |
| Qy       | 1686   | ATGCTTAAATAGCTTAAATATTTTACCCTGTACCTTTTTCACCAATCTATTATA     | 1745 |
| Db       | 1152   | ATGCTTAAATAGCTTAAATATTTTACCCTGTACCTTTTTCACCAATCTATTATA     | 1208 |
| Qy       | 1746   | ATATTTTTCATGAGCAAGTTCTCTCTGAG                              | 1774 |
| Db       | 1209   | ATATTTTTCATGAGCAAGTTCTCTCTGAG                              | 1237 |
| RESULT 8 |  |  |      |
| ID       | ABL90796   | Human polynucleotide standard; cDNA; 1249 BP.              |      |
| AC       | ABL90796;  |  |      |
| XX       | 24-MAY-2002 (first entry)  |  |      |
| DE       | Human polynucleotide SEQ ID NO 1358.                                     |  |      |
| KW       | Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;    |  |      |
| KW       | antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifolice;  |  |      |
| KW       | vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;    |  |      |
| KW       | cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder; |  |      |
| KW       | neurological disease; infection; human; secreted protein; gene; ss.      |  |      |
| XX       | Homo sapiens.  |  |      |
| OS       | WO200190304-A2.  |  |      |
| EN       | 29-NOV-2001.   |  |      |
| PD       | 18-MAY-2001; 2001WO-US16450.   |  |      |
| XX       | 19-MAY-2000; 2000US-205515P.   |  |      |
| PR       | (HUMA-) HUMAN GENOME SCT INC.  |  |      |
| PA       | Birse CE, Rosen CA;  |  |      |
| PI       | WPI; 2002-122018/16.   |  |      |
| DR       | P-PSDB; ABB90387.  |  |      |
| XX       | Novel 1405 isolated polypeptides, useful for diagnosis, treatment and    |  |      |
| PT       | prevention of neural, immune system, muscular, reproductive              |  |      |
| PT       | gastrointestinal, pulmonary, cardiovascular, renal and proliferative     |  |      |
| PT       | disorders -  |  |      |
| XX       | Claim 4; SEQ ID NO 1358; 2081bp + Sequence Listing; English.             |  |      |
| XX       | The invention relates to novel genes (ABL89449-ABL90853) and proteins    |  |      |
| CC       | (ABB9040-ABB9044) useful for preventing, treating or ameliorating        |  |      |
| CC       | medical conditions e.g. by protein or gene therapy. The genes are        |  |      |
| CC       | isolated from a range of human tissues disclosed in the specification.   |  |      |
| CC       | The nucleic acids, proteins, antibodies and (ant)agonists are useful     |  |      |
| CC       | in the diagnosis, treatment and prevention of: (a) cancer, e.g. Breast   |  |      |

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the CC printed specification, but was obtained in electronic format directly CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.  
 XX

Sequence 1249 BP; 379 A; 258 C; 273 G; 335 T; 4 other;

Query Match 65.8%; Score 1167.2; DB 24; Length 1249;  
 Best Local Similarity 98.7%; Pred. No. 1.3e-314;  
 Matches 1206; Conservative 1; Mismatches 11; Indels 4; Gaps 3;

QY 547 CTGTACCCCGACACGAATTTGACGGTCCGAGGTTGGGAGAGACCGTGAATCTTC 606  
 DB 18 CTGTACCCCGACACGAATTTGACGGTCCGAGGTTGGGAGAGACCGTGAATCTTC 77  
 QY 607 TTGTTGGCCCTCATCCACGCGCGGGGCTGACCTTCGACAGCAAGATCTGAC 666  
 DB 78 TTGTTGGCCCTCATCCACGCGCGGGGCTGACCTTCGACAGCAAGATCTGAC 137  
 QY 667 TTATGAGAGATGGGCGCTGAGAGCGGTGATCTTTAACTCCCGGAGACCGCAATGAGT 726  
 DB 138 TTATGAGAGAGGCGCGCTGAGAGCGGTGATCTTTAACTCCCGGAGACCGCAATGAGT 197  
 QY 727 CATCCCATGCTCACCCCGGAGTGAATGATGTTGCAATCATGATCGGCAATCTGAA 786  
 DB 198 CATCCCATGCTCACCCCGGAGTGAATGATGTTGCAATCATGATCGGCAATCTGAA 257  
 QY 787 AGGCACAAAATTCGATCATCTATCAAGAGCAATGAGTGAATGATGAGT 846  
 DB 258 AGGCACAAAATTCGATCATCTATCAAGAGCAATGAGTGAATGATGAGT 317  
 QY 847 AGGCAAAATTCGATCATCTATCAAGAGCAATGAGTGAATGATGAGT 906  
 DB 318 AGGCAAAATTCGATCATCTATCAAGAGCAATGAGTGAATGATGAGT 377  
 QY 907 TTTTATTTAAGCGCGGCACTGCGGCTATTTTAACTTTTCTGCTGAGAGCTGCG 966  
 DB 378 TTTTATTTAAGCGCGGCACTGCGGCTATTTTAACTTTTCTGCTGAGAGCTGCG 437  
 QY 967 GAATGCAAGAGCTCAAGAGCAAGAGCAATTTAAAGCAATGCTTAAAGCTAT 1026  
 DB 438 GAATGCAAGAGCTCAAGAGCAAGAGCAATTTAAAGCAATGCTTAAAGCTAT 497  
 QY 1027 TGAAGAGCTCAAGAGCAAGAGCAATTTAAAGCAATGCTTAAAGCTAT 1085  
 DB 498 TGAAGAGCTCAAGAGCAAGAGCAATTTAAAGCAATGCTTAAAGCTAT 557  
 QY 1086 ATAGTTGCTGCTGATGATGATTTAAAGCAATTTGATGATGATGAT 1145  
 DB 558 ATAGTTGCTGCTGATGATGATTTAAAGCAATTTGATGATGATGAT 617  
 QY 1146 GGAACCATATTTTCAATTAAGAGAGTGTGACCAAGGCTTAAACAGCAACTTGC 1205  
 DB 618 GGAACCATATTTTCAATTAAGAGAGTGTGACCAAGGCTTAAACAGCAACTTGC 677  
 QY 1206 CCAATGCAAAATGATGATTAAGAGTGTGAGTGAATGATGAT 1265  
 DB 678 CCAATGCAAAATGATGATTAAGAGTGTGAGTGAATGATGAT 737  
 QY 1266 CAGTCTTTTACAGAGCCCTGATCATTAATATCTATGCTCCCTCCCATGAG 1325  
 DB 738 CAGTCTTTTACAGAGCCCTGATCATTAATATCTATGCTCCCTCCCATGAG 797  
 QY 1326 AGGATATCGAGAGAGAGCGCATGATCTGATTTGCTTCACTAAGAGAGAGTAC 1385

DB 728 AGGATATCGAGAGAGAGCGCATCATCTGATTTGCTTCACTAAGAGAGTAC 857  
 QY 1386 CGCTCGAG 1445  
 DB 888 CGCTCGAG 917  
 QY 1446 CAAATTCGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1505  
 DB 918 CAAATTCGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 977  
 QY 1506 AAATTCGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1565  
 DB 978 AAATTCGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037  
 QY 1566 AAATTCGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1625  
 DB 1038 AAATTCGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097  
 QY 1626 TGAATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1685  
 DB 1098 TGAATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157  
 QY 1686 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1745  
 DB 1158 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1214  
 QY 1746 ATATTTTCAATGAGAGAGTTC 1767  
 DB 1215 ATATTTTCAATGAGAGAGTTC 1236

RESULT 9  
 AAT72094  
 ID AAT72094 standard; cDNA; 2145 BP.  
 XX  
 AC AAT72094;  
 XX  
 DT 25-MAR-2002 (first entry)  
 XX  
 DE Mouse GRAIL cDNA.  
 XX  
 KW Murine; human; GRAIL; anergy; attenuation; tyrosine; phosphorylation;  
 KW antigenic stimulation; interleukin-2; gene therapy; polymorphism; IL-2;  
 KW autoimmune disease; tumour cell; cancer; transplant rejection; ss.  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 358..1641  
 FT /\*tag= a  
 FT /product= "GRAIL"  
 XX  
 EN W0200165943-A1.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PE 11-MAY-2001; 2001W0-DS15385.  
 XX  
 PR 11-MAY-2000; 2000US-203513P.  
 XX  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 XX  
 FI Ford GS, Bloom D, Fatman CG;  
 XX  
 DR WPI; 2002-055597/07.  
 DR P-PDB; AAB47799.  
 XX  
 PT Novel nucleic acid sequences of anergy associated genes, including  
 PT GRAIL gene useful in the evaluation of pathophysiology or immunotherapy  
 PT of cancer, autoimmune disease and transplant rejection -  
 XX  
 PS Claim 2; Page 42-44; 50pp; English.  
 XX



XX AAC56340;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 519.  
XX  
KW Human; immunosuppressive; antiarthritic; ss; antineoplastic;  
KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin aging; food additive; preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
PN W0200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01341.  
XX  
PF 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234222.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235634.  
PR 27-SEP-2000; 2000US-0235636.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241121.  
PR 20-OCT-2000; 2000US-0241185.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250161.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251038.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-488783/53.  
 DR P-PSDB; AAU16353.  
 XX  
 XX New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 XX Claim 1; SEQ ID No 519; 980bp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. Antibodies to the proteins can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. Rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection, and many other  
 CC disorders listed in the specification. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence encodes a novel secreted protein of the invention.

Query Match 65.3%; Score 1157.6; DB 22; Length 1250;  
 Best Local Similarity 98.8%; Pred. No. 5e-312;  
 Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

QY 547 CTGTATCCCGACACGAAATTTCAACGCTGCCACGCTTGGGGAGACCGGTGCACTCTC 606  
 DB 18 CTGTATCCCGACACGAAATTTCAACGCTGCCACGCTTGGGGAGACCGGTGCACTCTC 77  
 QY 607 TTGGTTGGCCCTCATTCACACGCGGGGGGCTGACCTTGGAGCAAGATCATCTGGC 666  
 DB 78 TTGGTTGGCCCTCATTCACACGCGGGGGGCTGACCTTGGAGCAAGATCATCTGGC 137  
 QY 667 TTATAGAGATGGGCGCTGAGACCGGTATCTTAACTTCCCGGACCCGCAATAGGT 726  
 DB 138 TTATAGAGAGGGGCGCTGAGACCGGTATCTTAACTTCCCGGACCCGCAATAGGT 197  
 QY 727 CATCCCATGCTTACCCGGGTGCGATGAGATTTGGCAATCAATGATGGCAATCTGAA 786  
 DB 198 CATCCCATGCTTACCCGGGTGCGATGAGATTTGGCAATCAATGATGGCAATCTGAA 257

QY 787 AGGCACAAAATTTCTGATCTATTCGAAAGGCACTACAGTGAATGCTGATAGAGT 846  
 DB 258 AGGCACAAAATTTCTGATCTATTCGAAAGGCACTACAGTGAATGCTGATAGAGT 317  
 QY 847 AGGCAAAAACATGAGCCCTTGGGAGATCACTAACTCAA-TTTTTCCTGTTCTGTGCTCT 905  
 DB 318 AGGCAAAAACATGAGCCCTTGGGAGATCACTAACTCAA-TTTTTCCTGTTCTGTGCTCT 377  
 QY 906 TTTTATATATACGCGCGCAACTGTGGGCTATTTTATCTTTATCTGCTCGAAGGCTAC 965  
 DB 378 TTTTATATATACGCGCGCAACTGTGGGCTATTTTATCTTTATCTGCTCGAAGGCTAC 437  
 QY 966 GGAATGCAAGAGCTCAAGAGCAAGAGCAAGGCAATTTAAAGGCAAGTCAAAAAAGCTA 1025  
 DB 438 GGAATGCAAGAGCTCAAGAGCAAGAGCAAGGCAATTTAAAGGCAAGTCAAAAAAGCTA 497  
 QY 1026 TTGAAAGCTTCACTACGCACTGTAACAGAGACAAAG-AAATTTGCTGATGGA 1084  
 DB 498 TTGAAAGCTTCACTACGCACTGTAACAGAGACAAAGAAATTTGCTGATGGA 557  
 QY 1085 GATAGTTGCTGTGTCATGAAATTTGTAACCAATGATTTGGTACGATCTTAAG 1144  
 DB 558 GATAGTTGCTGTGTCATGAAATTTGTAACCAATGATTTGGTACGATCTTAAG 617  
 QY 1145 TGCACCAATTTTCCATTAAGACATGTTGACCCATGCTGTTAAACCAAGACTTGC 1204  
 DB 618 TGCACCAATTTTCCATTAAGACATGTTGACCCATGCTGTTAAACCAAGACTTGC 677  
 QY 1205 CCGATGCGAAATTTGACATCTCAAGGCTTTGGAGATTTGGATGTTGTAAGTGA 1264  
 DB 678 CCGATGCGAAATTTGACATCTCAAGGCTTTGGAGATTTGGATGTTGTAAGTGA 737  
 QY 1265 TCACTGCTTTTCAAGGCTCTGTTATTCATTAATTAATTAATTAATTAATTAATTA 1324  
 DB 738 TCACTGCTTTTCAAGGCTCTGTTATTCATTAATTAATTAATTAATTAATTAATTA 797  
 QY 1325 GAGATTAATGACAGAGACACGATCATCTGATGATGCTGATGACAGGACAGATGA 1384  
 DB 798 GAGATTAATGACAGAGACACGATCATCTGATGATGCTGATGACAGGACAGATGA 857  
 QY 1385 CCGGCTCTGAGAGACAGTGCATGCACTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1444  
 DB 858 CCGGCTCTGAGAGACAGTGCATGCACTCAAGTCAAGTCAAGTCAAGTCAAGTCA 917  
 QY 1445 GCAAAATCTGAGAGAGTGAATGTTATCCATGTTGACCAACCACTTGAAGAGAC 1504  
 DB 918 GCAAAATCTGAGAGAGTGAATGTTATTCCTCATGTTGACCAACCACTTGAAGAGAC 977  
 QY 1505 GAAACTCTTATCAAGAGACTGCTGTTGAGAAATTAATTTAAATCTGCTGAATAG 1564  
 DB 978 GAAACTCTTATCAAGAGACTGCTGTTGAGAAATTAATTTAAATCTGCTGAATAG 1037  
 QY 1565 AAAACTGAACCAATAGTAATTAACAGAACTGCAATCAAGGCTAGTTCTATTAATAA 1624  
 DB 1038 AAAACTGAACCAATAGTAATTAACAGAACTGCAATCAAGGCTAGTTCTATTAATAA 1097  
 QY 1625 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1684  
 DB 1098 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1157  
 QY 1685 TATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1744  
 DB 1158 TATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1214  
 QY 1745 AATAATTTTCAAGAGAGTTTC 1767  
 DB 1215 AATAATTTTCAAGAGAGTTTC 1237

RESULT 11  
 ABX73681  
 ID ABX73681 standard; DNA; 1250 BP.

XX AC ABX73681;  
 XX DT 18-MAR-2003 (first entry)  
 XX DE Human novel polynucleotide #509.  
 XX  
 KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002132753-A1.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 17-JAN-2001; 2001US-0764864.  
 XX  
 PR 31-JAN-2000; 2000US-179065P.  
 PR 04-FEB-2000; 2000US-180628P.  
 PR 28-JUN-2000; 2000US-214886P.  
 PR 07-JUL-2000; 2000US-216647P.  
 PR 07-JUL-2000; 2000US-216880P.  
 PR 11-JUL-2000; 2000US-217487P.  
 PR 11-JUL-2000; 2000US-217496P.  
 PR 14-JUL-2000; 2000US-218290P.  
 PR 26-JUL-2000; 2000US-220963P.  
 PR 14-AUG-2000; 2000US-224518P.  
 PR 14-AUG-2000; 2000US-224519P.  
 PR 14-AUG-2000; 2000US-225267P.  
 PR 14-AUG-2000; 2000US-225268P.  
 PR 14-AUG-2000; 2000US-225270P.  
 PR 14-AUG-2000; 2000US-225447P.  
 PR 14-AUG-2000; 2000US-225757P.  
 PR 14-AUG-2000; 2000US-225758P.  
 PR 22-AUG-2000; 2000US-226868P.  
 PR 30-AUG-2000; 2000US-228924P.  
 PR 01-SEP-2000; 2000US-229287P.  
 PR 01-SEP-2000; 2000US-229343P.  
 PR 01-SEP-2000; 2000US-229344P.  
 PR 01-SEP-2000; 2000US-229345P.  
 PR 05-SEP-2000; 2000US-229509P.  
 PR 05-SEP-2000; 2000US-229513P.  
 PR 21-SEP-2000; 2000US-234223P.  
 PR 21-SEP-2000; 2000US-234274P.  
 PR 25-SEP-2000; 2000US-234997P.  
 PR 27-SEP-2000; 2000US-235834P.  
 PR 29-SEP-2000; 2000US-236327P.  
 PR 29-SEP-2000; 2000US-236367P.  
 PR 29-SEP-2000; 2000US-236368P.  
 PR 29-SEP-2000; 2000US-236369P.  
 PR 02-OCT-2000; 2000US-236802P.  
 PR 02-OCT-2000; 2000US-237037P.  
 PR 02-OCT-2000; 2000US-237038P.  
 PR 02-OCT-2000; 2000US-237039P.  
 PR 02-OCT-2000; 2000US-237040P.  
 PR 13-OCT-2000; 2000US-239935P.  
 PR 20-OCT-2000; 2000US-240960P.  
 PR 20-OCT-2000; 2000US-241785P.  
 PR 20-OCT-2000; 2000US-241809P.  
 PR 01-NOV-2000; 2000US-244617P.  
 PR 17-NOV-2000; 2000US-248999P.  
 PR 08-DEC-2000; 2000US-251856P.  
 PR 08-DEC-2000; 2000US-251868P.  
 PR 08-DEC-2000; 2000US-251869P.

XX (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 XX Rosen Ch, Ruben SM, Barash SC;  
 XX  
 XX WPI: 2003-147444/14.  
 DR P-PSDB; ABUS5421.  
 XX  
 PT New polypeptides and nucleic acids, useful in gene therapy for  
 PT treating, inhibiting or preventing e.g. neural immune system,  
 PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,  
 PT cardiovascular or renal disorders -  
 XX  
 XX  
 PS Claim 1; SEQ ID NO 519; 402bp; English.  
 XX  
 CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid  
 CC arthritis and multiple sclerosis), muscular disorders, respiratory  
 CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),  
 CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,  
 CC cardiovascular disorders (e.g. congenital heart defects, Epstein's  
 CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute  
 CC kidney failure and end-stage renal disease), hyperproliferative disorders  
 CC (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g.  
 CC septic shock, bursitis and appendicitis), allergic reactions and  
 CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,  
 CC atherosclerosis and myocardial infarction) and cancerous diseases.  
 CC Sequences ABX73173-ABX74167 represent human novel polynucleotides of the  
 CC invention.  
 XX  
 SQ Sequence 1250 BP; 380 A; 258 C; 274 G; 336 T; 2 other;  
 XX  
 Query Match 65.3%; Score 1157.6; DB 25; Length 1250;  
 Best Local Similarity 98.8%; Pred. No. 5e-312;  
 Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;  
 XX  
 QY 547 CTGTAAACCCGACAGAAATTTTACGGTGGCCACGGTTTGGGAGACACCGTGAAGTTCG 606  
 DB CTGTAAACCCGACAGAAATTTTACGGTGGCCACGGTTTGGGAGACACCGTGAAGTTCG 77  
 QY 607 TTGTGTGGCCCTCATCAACAGCGGGGGGCTGCACCTTCGAGCAAGATCCATCTGAC 666  
 DB TTGTGTGGCCCTCATCAACAGCGGGGGGCTGCACCTTCGAGCAAGATCCATCTGAC 137  
 QY 727 CATCCCAATGCTCACCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 786  
 DB CATCCCAATGCTCACCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 197  
 QY 727 CATCCCAATGCTCACCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 786  
 DB CATCCCAATGCTCACCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 197  
 QY 787 AGGCAAAAATTTCTCAATCTATTCAAGAGGCAATGATGATGATGATGATGATGATGATGAT 846  
 DB AGGCAAAAATTTCTCAATCTATTCAAGAGGCAATGATGATGATGATGATGATGATGATGAT 317  
 QY 847 AGGCAAAAATTTCTCAATCTATTCAAGAGGCAATGATGATGATGATGATGATGATGATGAT 905  
 DB AGGCAAAAATTTCTCAATCTATTCAAGAGGCAATGATGATGATGATGATGATGATGATGAT 377  
 QY 906 TTTTATATTAACGGGGGCACTGTGGGCTATTATTTCTTTATCTGCTGAGAGGCTTAC 965  
 DB TTTTATATTAACGGGGGCACTGTGGGCTATTATTTCTTTATCTGCTGAGAGGCTTAC 437  
 QY 966 GGAATGCAAGAGCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1025  
 DB GGAATGCAAGAGCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 497  
 QY 1026 TTGGAAGGCTTCAATGAGCACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084







|   |      |   |      |
|---|------|---|------|
| QY  | 683  | TCGAGACGGCATCTTAACTCCCGGAGCCCGCATGAGGCTATCCCATGTCTAC            | 742  |
| Db  | 421  | TCGAGAGGGTATTTTAACTCCCTGGAGCCCGCATGAGGCTATCCCATGTCTAC           | 480  |
| QY  | 743  | CCGGGTGACATGATCTGTTGCAATCATGATCGCAATCGAAAGGACCAAAATCTTG         | 802  |
| Db  | 481  | CCGGGTGAGGGGACATGTTGCAATTCATTCATTCGCAATCGAAAGGACCAAAATCTTG      | 540  |
| QY  | 803  | CAATCTATTCAAAGAGCATTAAGTGAATGTGTCAATGATAGAGAGGAAAAACATGGC       | 862  |
| Db  | 541  | CAGTATTCATTAAGAGGCATTAAGTGAATGTGTCAATGATAGAGAGGAAAAACATGGC      | 600  |
| QY  | 863  | CCTTGGGTAACTATTCATTTTCTGTTCTGTGTCTTTTATTAATTACGGCG              | 922  |
| Db  | 601  | CCTTGGGTAACTATTCATTTTCTGTTCTGTGTCTTTTATTAATTACGGCA              | 660  |
| QY  | 923  | GCAACGTGGGCTATTATTCCTTTATCTGTCTGCAAGGCTGATGGAATGCAAGAGCTCA      | 982  |
| Db  | 661  | GCAACGTGGGCTATTATTCCTTTATCTGTCTGCAAGGCTGATGGAATGCAAGAGCTCA      | 720  |
| QY  | 983  | AGCAGGAGACAGAGGCTATTTAAAGGACATGCTAAAAAGCTATTTGAAAGGCTTACA       | 1042 |
| Db  | 721  | AGTAGGAGACAGAGGCTATTTAAAGGACATGCTAAAAAGGCTATTTGAAAGGCTTACA      | 780  |
| QY  | 1043 | CGCACCTGTAACCAAGAGCAAGGAATTTGGCCTGATGAGATGTTGTGCTGTGG           | 1102 |
| Db  | 781  | CGCACCTGTAACCTGAGACCAAGGAATTTGGCCTGATGAGATGTTGTGCTGTGG          | 840  |
| QY  | 1103 | ATTGAATGTATTAACCAATATGATTTGGATGCGATCTTAAGTGTCAACATATTTTCCAT     | 1162 |
| Db  | 841  | ATTGACCTATTAAGCCCAATATGATTTGGATGCGATCTTAAGTGTCAACATATTTTCCAT    | 900  |
| QY  | 1163 | AAGCACTGATGACCACTGGCTGTTAAACAACAACATCTGCCCATGTGCAGAAATGTAC      | 1222 |
| Db  | 901  | AAGCACTGTGTGACCCGTGATGTTTAAACAACAAGACCTGCCCATGTGTGACAG          | 960  |
| QY  | 1223 | ATATCTAAGCTTTGGGAAATTGAGTGTGATGTGAAGATGATGATGATCTTTACAGTTC      | 1282 |
| Db  | 961  | ATATCTAAGCTCTGGGAAATTGAGTGTGATGTGAAGATGATGATGATCTTTACAGTTC      | 1020 |
| QY  | 1283 | CCGTGATCCCAATGAATATATATATAGTGTGCTCCCTCCCATGAAGAGATTAATGACAGAG   | 1342 |
| Db  | 1021 | CCGTGATCCCAATGAAGATATATATATAGTGTGCTCCCTCCCATGAAGAGATTAATGACAGAG | 1080 |
| QY  | 1343 | ACGCGATCATCTGATATGCTTGAGTACAGGGAACATGAACCGCTCTTGGAGAAC          | 1402 |
| Db  | 1081 | ACGCGATCATCTGATATGCTTGAGTACAGGGAACATGAAGCGCTCTTGGAGAACAT        | 1140 |
| QY  | 1403 | GTGCACTGCAAGAAATGATCTACAGCTGTGTAACATGAAAGCAATCTCTGGCAGTG        | 1462 |
| Db  | 1141 | GTGCACTGCAAGAAATGATCTACAGCTGTGTAACATGAAAGCAATCTCTGGCAGTG        | 1200 |
| QY  | 1463 | GATGTTATTCCTCATGATGACCAACCACTTTGAAAGAGACGAATCTCTATCTAAGAG       | 1522 |
| Db  | 1201 | GATGTTATTCCTCATGATGACCAACCACTTTGAAAGAGAGCAATCTCTATCTAAGAG       | 1260 |
| QY  | 1523 | ACTGCTGTTCGAGAAATTAATCTTAA                                      | 1549 |
| Db  | 1261 | GCGCTGTGTTGGGATTAATCTTAA  | 1287 |
| RESULT 14   |      |   |      |
| ABK12990 standard; DNA; 1287 BP.                                  |      |   |      |
| ID  |      |   |      |
| ABK12990;   |      |   |      |
| AC  |      |   |      |
| XX  |      |   |      |
| DT  |      |   |      |
| XX  |      |   |      |
| 23-APR-2002 (first entry)   |      |   |      |
| DB  |      |   |      |
| Synthetic goliatsh DNA sequence #4.                               |      |   |      |
| Goliatsh protein; antiangiogenic; vasotropic; gene therapy;       |      |   |      |
| KW  |      |   |      |
| dosage form; angiogenesis; neurogenesis; tumour; vascularisation; |      |   |      |

KM cancer; ischemia; neuroblastoma; neurodegenerative disease;  
KW unrecovered nerve trauma; ds.  
XX Synthetic.  
XX  
XX W0200193681-A1.  
XX  
XX 13-DEC-2001.  
XX  
XX 01-JUN-2001; 2001MO-US18000.  
XX  
XX 02-JUN-2000; 2000US-0586398.  
XX  
XX (REGC) UNIV CALIFORNIA.  
XX  
XX Harland R, Baker JC;  
XX  
XX WPI, 2002-147637/19.  
XX  
XX New compositions comprising gollioth proteins, useful for modulating  
XX angiogenesis or neurogenesis in mammals e.g. for preventing or  
XX treating undesirable vascularrisation of a tumour, ischaemia or  
XX neurodegenerative disease -  
XX  
XX Disclosure, Page 41; 45pp; English.

The present invention relates to a new pharmaceutical composition that comprises a gollioth polypeptide in dosage form. The gollioth polypeptide has a sequence identity of at least 75% to the protein sequences (AAU47418-AAU47421) fully defined in the specification. The composition is useful for modulating angiogenesis or neurogenesis in mammals, particularly in humans or mice. Specifically, the composition is useful for the prophylactic and/or therapeutic treatment of excess angiogenesis e.g. undesirable vascularrisation of a tumour or insufficient angiogenesis e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma. The present nucleic acid sequence represents synthetic DNA sequence #4 that is one of several artificial gollioth nucleic acids (ABK12987-ABK12992) used in the invention for modulating angiogenesis or neurogenesis.

Sequence 1287 BP; 312 A; 321 C; 356 G; 298 T; 0 other;

|    | Query Match   | 58.8%          | Score 1042.8       | DB 24    | Length 1287 |  |
|----|---|----------------|--------------------|----------|-------------|--|
|    | Best Local Similarity   | 88.1%          | Pred. No. 5.3e-280 |          |             |  |
|    | Matches 1134  | Conservative 0 | Mismatches 153     | Indels 0 | Gaps 0      |  |
| QY | 263 ATGGGCGCGCCGCTTGTGGGCGCGGAGTCCTGTGCACGCGTGATGCTGGCGCTTTTCAGATTG | 322            |                    |          |             |  |
| Db | 1 ATGGGCGCGCGCCGCGGATAAGGAGTCACTGCCGCGGTGTGCTGGAGAGCTGCCCGCCCTA     | 60             |                    |          |             |  |
| QY | 323 CTGGGATGATGGCTTCCGTGCTGCACCCTGAATCCGAGGACAACGGATCCCGGGGGGCTGAA  | 382            |                    |          |             |  |
| Db | 61 CTGGCTTGTATGCTTCTTCTGTGCTTGAATCATACGCCCCCGGTCCCGTAGAGCGGA        | 120            |                    |          |             |  |
| QY | 383 GCAATGTGAACCGCGTAATCAAAGTCTCTGCGCGGTTTCGACACGGAGTGAACCGT        | 442            |                    |          |             |  |
| Db | 121 GCCCTGTGACACCGGTAATCAACTGTCTCTGCGCGGTTTCGACAAACCGAGTGAACCGC     | 180            |                    |          |             |  |
| QY | 443 AGCGATGAGGAGCTGACGACAGAGAGGCGGTGAACGAGCACAGCTGCACGCTGAGAGCTGTG  | 502            |                    |          |             |  |
| Db | 181 ACTGTGTGAGAGCTGACGACAGAGAGGCGGTGAACGAGCACAGCTGCACGCTGAGAGCTGTG  | 240            |                    |          |             |  |
| QY | 503 GCCTGGGAGCTCTGTGAACCGCCGACGGGCTCCGGGGCGCTTAAAGCTGTAAACCGGACAAG  | 562            |                    |          |             |  |
| Db | 241 TCCTGGGAGCTCTGTGTTCGCGCCACAGGGCCCGGCGCCTAAAGCTGTAAACCGGACAAC    | 300            |                    |          |             |  |
| QY | 563 AATTTCACGAGCCACGAGTTTGGGAGAAGCAACGTGCAAGTCTCTGTGGTTGGCCCTCATC   | 622            |                    |          |             |  |
| Db | 301 AATTTCACAGTCCACAGTTTGGGAGTGAACGAGTGAAGTATCTGGTTGGCCCTCATC       | 360            |                    |          |             |  |
| QY | 623 CAACGCGCGGAGGCTTGCATTCCTGCAAGAAATCAATCTTATGAGAAATGGGG           | 682            |                    |          |             |  |

|           |                                    |  |      |
|-----------|------------------------------------|--|------|
| Db        | 361                                | CAGCGCGGTGGAAGGCTGCAACATGCGCGGACAGAGATCAATCGGCTTCAGAGAGAGGCGCT     | 420  |
| Qy        | 683                                | TTTGGAGGCCGTCACTCTTTAAGCTTCCCGGGACCCGCAATGAGTGCATCCCATGTTCTCAC     | 742  |
| Db        | 421                                | TCTGAGACCGGTGTCAATGTTTAACTTCCCTGGGACACGCAATGAGGTCAATCCCTATGTTCTCAC | 480  |
| Qy        | 743                                | CCGGGTGCAGTGAACATTTGTGCAATCAATGATCGGCAATCTGAAAGGCAAAAAATTCG        | 802  |
| Db        | 481                                | CCGGGTGCCGGGACCTTGTTCGATGATGATTTGGCAATCTGAAAGGACAAAAATTCG          | 540  |
| Qy        | 803                                | CAATCAATTCGAAAGGCGATGCAAGTGAACATGTCATPAGAAATGAGGAAAAACATGAGC       | 862  |
| Db        | 541                                | CATTCATTCGAAAGAGGCAATCCAGTCAATATGTCATGAGAAATGAGGAAAAACAGAGC        | 600  |
| Qy        | 863                                | CCCTGGGAGATACAAATCAATTTTTTGTTCGATCGATCCCTTTTATATTAATPACGGCG        | 922  |
| Db        | 601                                | CCCTGGGTGATCATTTATTCATTTTTCTGTGCTCTGTGTCTCCTTTTTCATGATACGGCA       | 660  |
| Qy        | 923                                | GCAACTGTGCGATTAATTTATCTTTTATTCCTGTGTGAGAGCTACGAAATGCAGAGCTCA       | 982  |
| Db        | 661                                | GCAACCGATAGGCTATTTACTCTTTATTCCTGTGTGAGAGTATCCCAATGCAAGAGCTCA       | 720  |
| Qy        | 983                                | AGCAGGAAGCAGAGGCAATTTAAAGGAGATGCTPAAAGAAAGCATTTGAGAGGCTTCACTA      | 1042 |
| Db        | 721                                | AGGAGGAAGCAGAGGCAATTTAAAGGCAATGCTAATAAAATGCAATTTGAAAGCTTCACTG      | 780  |
| Qy        | 1043                               | GCGCACTGTAACAGAGGACAGAGGAATTTGGCTCTGATGAGATAGTTGTGTGTGTGC          | 1102 |
| Db        | 781                                | GCGACCTTGAAACAGAGGACAGAGGAATTTGACCTGAGAGATACCTGTGTGTGTGC           | 840  |
| Qy        | 1103                               | ATTGAATTTGATTAACCAATATGTTGGTGAAGCATTTAGAGTCAACATATTTCCAT           | 1162 |
| Db        | 841                                | ATTGAGCTCATTAAGCAAAATGATTTGGTGGAGATCTCACTGAATCATATTTCCAT           | 900  |
| Qy        | 1163                               | AAGACATGTGTGAACCCATGCGCTGTAAACAGACACCTGCCCATGTGCAAAATGTGAC         | 1222 |
| Db        | 901                                | AATTAATGTGTGACCCGCGCTTTTGAACAGAGACGCTGCCCATGTGCAAAATGTGAC          | 960  |
| Qy        | 1223                               | ATACTCAAAGCTTTGGGAATTTGAGGTGATGTGTGAAGATGATCAATGCTTTACAGTC         | 1282 |
| Db        | 961                                | ATTCTCAAAGCTTTGGGAATTTGAGGTGATGTGATGATGATGATCAATGCTTTACAGTT        | 1020 |
| Qy        | 1283                               | CCGTGATCCCAATTAATATCTAATTAATGTGCGCTCCCTCCCATGAAAGAGATATGCGACGAG    | 1342 |
| Db        | 1021                               | CCGTGTTTAATTAATTAATCAATCTAATATCTGCTCTCCCATGAAAGAGACATGTGCGAGTG     | 1080 |
| Qy        | 1343                               | ACCGCATCAATCTGATATATCTTCAAGTACAGGAGAACAGATGACCGGCTCTGAGAGACAC      | 1402 |
| Db        | 1081                               | ACCGCATCAATCTGATATATCTATCAAGTACAGGAGACAGATGAGACCACTCTGAGAGACAT     | 1140 |
| Qy        | 1403                               | GTGCAGTCAACAATAAGTCTAAGCTGTGTTAAACATGAAAGCAAAATTCGTGTGAGTG         | 1462 |
| Db        | 1141                               | GCGCAGTCAACAAGAAATCTTACAGTGTGTAAACATGAAAGCAAAATTCGTGTGAGTG         | 1200 |
| Qy        | 1463                               | GATGTATTTCTCATATGTCAGAACCCACCTTTGAAAGAGACGAAATCTCTTAATCAAG         | 1522 |
| Db        | 1201                               | GATGTATTTCTCATATGTCAGAACCCACCTTTGAAAGAGATGAAATCTCTGTATCAAG         | 1260 |
| Qy        | 1523                               | ACTGCTGTTGAGAAATTAATCTTAA  | 1549 |
| Db        | 1261                               | GCTGCTGTTGGAGATTAATCTTTAA  | 1287 |
| RESULT 15 |                                    |  |      |
| ABK12991  |                                    |  |      |
| ID        | ABK12991                           | standard; DNA; 1287 bp.  |      |
| XX        | ABK12991;                          |  |      |
| AC        |                                    |  |      |
| XX        | 23-APR-2002                        | (first entry)  |      |
| XX        |                                    |  |      |
| XX        | Synthetic goliath DNA sequence #5. |  |      |
| XX        |                                    |  |      |

|    |  |
|----|--|
| KM | Goliath protein; antiangiogenic; vasotropic; gene therapy;                       |
| KN | dosage form; angiogenesis; neurogenesis; tumour; vascularisation;                |
| KX | cancer; ischaemia; neuroblastoma; neurodegenerative disease;                     |
| KY | unrecovered nerve trauma; ds.  |
| XX | Synthetic.   |
| OS |  |
| XS | WO200193681-A1.  |
| PN |  |
| PD | 13-DEC-2001.   |
| PP | 01-JUN-2001; 2001MO-US18000.   |
| XX |  |
| FR | 02-JUN-2000; 2000US-0586398.   |
| PA | (REGC ) UNIV CALIFORNIA.   |
| PI |  |
| DR | Hartland R, Baker JC;  |
| XX |  |
| XX | WPI, 2002-147637/19.   |
| PT |  |
| PT | New compositions comprising goliath proteins, useful for modulating              |
| PT | angiogenesis or neurogenesis in mammals e.g. for preventing or                   |
| PT | treating undesirable vascularisation of a tumour, ischaemia or                   |
| PS | neurodegenerative disease  |
| PS | -  |
| PS | Disclosure; Page 41-42; 45pp; English.   |
| XX |  |
| CC | The present invention relates to a new pharmaceutical composition that           |
| CC | comprises a goliath polypeptide in dosage form. The goliath polypeptide          |
| CC | has a sequence identity of at least 75% to the protein sequences                 |
| CC | (AAU04918-AAU04921) fully defined in the specification. The composition          |
| CC | is useful for modulating angiogenesis or neurogenesis in mammals,                |
| CC | particularly in humans or mice. Specifically, the composition is useful          |
| CC | for the prophylactic and/or therapeutic treatment of excess angiogenesis         |
| CC | e.g. undesirable vascularisation of a tumour or insufficient angiogenesis        |
| CC | e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient           |
| CC | neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.         |
| CC | The present nucleic acid sequence represents synthetic DNA sequence #5           |
| CC | that is one of several artificial goliath nucleic acids (ARK12987-               |
| CC | ARK13992) used in the invention for modulating angiogenesis or                   |
| CC | neurogenesis.  |
| CC |  |
| SQ | Sequence 1287 BP; 312 A; 321 C; 356 G; 298 T; 0 other;                           |
|    |  |
|    | Query Match            58.8%; Score 1042.8; DB 24; Length 1287;                  |
|    | Best Local Similarity   88.1%; Pred. No. 5.3e-280;                               |
|    | Matches 114; Conservative   0; Mismatches 153; Indels         0; Gaps         0; |
| OY | 263 AAGGGGCGCGCGCTGGGCGCGGGGTCCGCGCGCGGTGGTGGCTTTCAATGG 322                      |
| DB | 1 AATGGGGCGCGCGCGCGGATTGGGGTCTACTGCCGCGCGGTGGAGACTGCCACATA 60                    |
| OY | 323 CTGCATGATGTCCTCTTGCTGACCCTGAATGCSGACGACCCGATTCCCGGGGGCTGAA 382               |
| DB | 61 CTGGCTTGAGCTTGCTTCTTGACTCTAGTCTCAAGCGCCCGCGTTCGCGAGCGGA 120                   |
| OY | 383 GAAGTGGAGACCCGTAACCTAAGTGTCCGTGGCGGCTTCCGACAAGGAGTGAACGAT 442                |
| DB | 121 GCCGTGTAACCGCTACTCACTGCTCTGGCGGGTTCGCAATCCGAGATGAACCGC 180                   |
| OY | 443 AGGTGGAGACTTAGCGCAGAAGGCGTGAACGSCAGACTGCGCTGAGCCTTG 502                      |
| DB | 181 ACCGTGGAGAGCTGAGCANAAAGGGGTGTAAAGGCAAGATCGCGCTGAGACTGTGC 240                 |
| OY | 503 GCTGGGCTCTGATGTAACCGCGCACGSGCGCGGGGACCTTAACGACTGTAACCGGACAAG 562             |
| DB | 241 TCCGGGGTCTGGATCCCGCGCGAGCGGCCGAGGCGCTTAAGCTGTAAAGCGGACCAAC 300               |
| OY | 563 AATTTCACGATGCCACGATTTGGGGAAGAACCGTGAAGTCTGTGGTTGGCCCTCATC 622                |
| DB | 301 AATTTCATGTGCCCACGATTTGGGGAAGAACGATGTAAGTATGATGTGGCCCTCATC 360                |



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 16:32:57, Search time 461 Seconds  
(without alignments)  
9816.120 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774  
Sequence: 1 agctgagctccaccgcggt.....atagcaagcttcctcctcag 1774

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PC1\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                                |
|------------|--------|-------------|--------|-------|--|
| 1          | 1178.8 | 66.4        | 2306   | 10    | US-09-764-864-63<br>Sequence 63, Appl      |
| 2          | 1157.6 | 65.3        | 1250   | 10    | US-09-764-864-519<br>Sequence 519, Appl    |
| 3          | 372.6  | 21.0        | 643    | 10    | US-09-878-178-1815<br>Sequence 1815, Appl  |
| 4          | 372.6  | 21.0        | 643    | 13    | US-10-046-935-1815<br>Sequence 1815, Appl  |
| 5          | 372.6  | 21.0        | 643    | 14    | US-10-146-502-1815<br>Sequence 1815, Appl  |
| 6          | 367.4  | 20.7        | 509    | 14    | US-10-066-543-1896<br>Sequence 1896, Appl  |
| 7          | 320.6  | 18.1        | 1188   | 12    | US-09-814-351-21198<br>Sequence 21198, A   |
| C 8        | 310    | 17.5        | 435    | 10    | US-09-920-300A-1176<br>Sequence 1176, Appl |
| C 9        | 310    | 17.5        | 435    | 12    | US-10-033-528-1176<br>Sequence 1176, Appl  |
| C 10       | 274.6  | 15.5        | 387    | 10    | US-09-983-965-5551<br>Sequence 5551, Appl  |
| 11         | 224.6  | 12.7        | 616    | 14    | US-10-060-036-25<br>Sequence 25, Appl      |
| 12         | 224.6  | 12.7        | 374    | 10    | US-09-983-965-4662<br>Sequence 4662, Appl  |
| 13         | 214.4  | 12.1        | 1200   | 12    | US-10-074-511-3<br>Sequence 3, Appl        |
| 14         | 209    | 11.8        | 2063   | 9     | US-09-935-390A-1<br>Sequence 1, Appl       |
| 15         | 209    | 11.8        | 392    | 10    | US-09-983-965-5467<br>Sequence 5467, Appl  |
| 16         | 201.2  | 11.3        |        |       |  |

|      |       |      |      |    |   |
|------|-------|------|------|----|---|
| 17   | 187.4 | 10.6 | 340  | 10 | US-09-783-590-3523<br>Sequence 3523, Appl |
| 18   | 180.8 | 10.2 | 3615 | 11 | US-09-822-846-596<br>Sequence 596, Appl   |
| 19   | 149.2 | 8.4  | 1322 | 9  | US-09-935-350A-16<br>Sequence 16, Appl    |
| 20   | 147.6 | 8.3  | 1253 | 14 | US-10-268-473-6<br>Sequence 6, Appl       |
| C 21 | 144.6 | 8.2  | 1251 | 12 | US-09-814-353-17086<br>Sequence 17086, A  |
| 22   | 144.4 | 8.1  | 1251 | 12 | US-09-933-767-156<br>Sequence 156, Appl   |
| 23   | 144.4 | 8.1  | 1251 | 14 | US-10-023-282-156<br>Sequence 156, Appl   |
| C 24 | 136.4 | 7.7  | 442  | 12 | US-09-814-353-4398<br>Sequence 4398, Appl |
| C 25 | 136.4 | 7.7  | 442  | 12 | US-09-814-353-10702<br>Sequence 10702, A  |
| 26   | 133.6 | 7.5  | 619  | 10 | US-09-764-864-485<br>Sequence 485, Appl   |
| 27   | 132   | 7.4  | 508  | 11 | US-09-918-995-21070<br>Sequence 21070, A  |
| C 28 | 117.2 | 6.6  | 463  | 9  | US-09-864-761-11364<br>Sequence 11364, A  |
| 29   | 112.8 | 6.4  | 826  | 9  | US-09-822-849A-90<br>Sequence 90, Appl    |
| C 30 | 112.4 | 6.3  | 435  | 10 | US-09-983-965-5079<br>Sequence 5079, Appl |
| 31   | 105.6 | 6.0  | 382  | 9  | US-09-815-343-1328<br>Sequence 1328, Appl |
| 32   | 105.2 | 5.9  | 253  | 9  | US-09-864-751-27944<br>Sequence 27944, A  |
| 33   | 102.8 | 5.8  | 1074 | 12 | US-09-971-382-211<br>Sequence 211, Appl   |
| 34   | 90.4  | 5.1  | 2330 | 11 | US-09-822-846-240<br>Sequence 240, Appl   |
| 35   | 86.4  | 4.9  | 922  | 9  | US-09-800-723-23<br>Sequence 23, Appl     |
| 36   | 80    | 4.5  | 1247 | 14 | US-10-106-698-2093<br>Sequence 2093, Appl |
| 37   | 79.4  | 4.5  | 549  | 11 | US-09-764-891-126<br>Sequence 126, Appl   |
| 38   | 79.2  | 4.5  | 455  | 14 | US-10-106-698-1892<br>Sequence 1892, Appl |
| 39   | 78.8  | 4.4  | 570  | 10 | US-09-734-017A-7<br>Sequence 7, Appl      |
| 40   | 78.6  | 4.4  | 439  | 9  | US-09-925-289-484<br>Sequence 484, Appl   |
| 41   | 78.6  | 4.4  | 439  | 11 | US-09-925-289-484<br>Sequence 484, Appl   |
| 42   | 78.4  | 4.4  | 583  | 12 | US-10-160-162-74<br>Sequence 74, Appl     |
| 43   | 78.4  | 4.4  | 1701 | 12 | US-09-930-213-312<br>Sequence 312, Appl   |
| 44   | 78.4  | 4.4  | 2346 | 12 | US-10-011-200-2<br>Sequence 2, Appl       |
| 45   | 78    | 4.4  | 726  | 11 | US-09-986-480-36<br>Sequence 36, Appl     |

#### ALIGNMENTS

RESULT 1  
US-09-764-864-63  
; Sequence 63, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rozen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antipodotes  
; FILE REFERENCE: P1223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; PRIOR APPLICATION DATE: 2001-01-17  
; Prior application data removed - consult PAM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 63  
; LENGTH: 2306  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-764-864-63

Query Match 66.4%; Score 1178.8; DB 10; Length 2306;  
Best local similarity 99.0%; Pred. No. 2.4e-314;  
Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

|    |     |   |     |
|----|-----|---|-----|
| QY | 547 | CTGTAACCGGACACAAATTTCACGGTCCGACGCTTTGGGAAAGCACTGCAAGTTC | 606 |
| DB | 12  | CTGTAACCGGACACAAATTTCACGGTCCGACGCTTTGGGAAAGCACTGCAAGTTC | 71  |
| QY | 607 | TTGGTGGCCCTCATCCAGACGCGCGGGGCTGCACTTCGCAAGCAATCATCTGGC  | 666 |
| DB | 72  | TTGGTGGCCCTCATCCAGACGCGCGGGGCTGCACTTCGCAAGCAATCATCTGGC  | 131 |
| QY | 667 | TTATGAGAGTGGGGCTGAGAGCCCTCATCTTAACTCCCGGAAACCGGAAAGAGT  | 726 |
| DB | 132 | TTATGAGAGTGGGGCTGAGAGCCCTCATCTTAACTCCCGGAAACCGGAAAGAGT  | 191 |
| QY | 727 | CATCCCAATCTCCACCCGGGAGTGAAGCAATTTGCAATCATGATGGCAATCTGA  | 786 |
| DB | 192 | CATCCCAATCTCCACCCGGGAGTGAAGCAATTTGCAATCATGATGGCAATCTGA  | 251 |

QY 787 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTCAAGT 846  
 DB 252 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTCAAGT 311  
 QY 847 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTCAAGT 905  
 DB 312 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTCAAGT 371  
 QY 906 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTTATCTTTATCTGTCTCGAAGGCTAC 965  
 DB 372 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTTATCTTTATCTGTCTCGAAGGCTAC 431  
 QY 966 GGAATGCAAGGCTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 1025  
 DB 432 GGAATGCAAGGCTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 491  
 QY 1026 TTGGAGAGGCTTCACTACGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1085  
 DB 492 TTGGAGAGGCTTCACTACGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 551  
 QY 1086 ATAGTGTGCTGTGTCATGTGATTTGATTAACCAATGATTTGTAGCATCTTAAGT 1145  
 DB 552 ATAGTGTGCTGTGTCATGTGATTTGATTAACCAATGATTTGTAGCATCTTAAGT 611  
 QY 1146 GCAACCAATTTTCCATTAAGACATGTGTGACCATGCTGTGTTAAACAGAGCTTGC 1205  
 DB 612 GCAACCAATTTTCCATTAAGACATGTGTGACCATGCTGTGTTAAACAGAGCTTGC 671  
 QY 1206 CCATGTCATATGTGCACTACTCAAGCTTTGGGAAATGAGTGAGATGTTGAAGAT 1265  
 DB 672 CCATGTCATATGTGCACTACTCAAGCTTTGGGAAATGAGTGAGATGTTGAAGAT 731  
 QY 1266 CAGTGTCTTTACAGTCCCTGTATCCATGATTAATCTTAATAGTCCCTCCCAATGAG 1325  
 DB 732 CAGTGTCTTTACAGTCCCTGTATCCATGATTAATCTTAATAGTCCCTCCCAATGAG 791  
 QY 1326 AGGATATGCGACGAG 1385  
 DB 792 AGGATATGCGACGAG 851  
 QY 1386 CGCCTCTGAGAGAGACGTCGACGTCAAAATGAAAGTCAAGCTGTGTAACCAATGAG 1445  
 DB 852 CGCCTCTGAGAGAGACGTCGACGTCAAAATGAAAGTCAAGCTGTGTAACCAATGAG 911  
 QY 1446 CAATGTCGTGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1505  
 DB 912 CAATGTCGTGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 971  
 QY 1506 AAACCTCAATCAAG 1565  
 DB 972 AAACCTCAATCAAG 1031  
 QY 1566 AAACCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1625  
 DB 1032 AAACCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1091  
 QY 1626 TGGATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1685  
 DB 1092 TGGATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1151  
 QY 1686 ATAGTAAATAGCTTAATAATTTTAATCTTTAATCTTTTCCACCAATCAATTA 1745  
 DB 1152 ATAGTAAATAGCTTAATAATTTTAATCTTTAATCTTTTCCACCAATCAATTA 1208  
 QY 1746 ATATTTTTCATAGGCAAGTTTCTCTCAG 1774  
 DB 1209 ATATTTTTCATAGGCAAGTTTCTCTCAG 1237

RESULT 2  
 US-09-764-864-519  
 ; Sequence 519, Application US/09764864  
 ; Patent No. US0020132753A1

GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT23  
 CURRENT APPLICATION NUMBER: US/09/764,864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ. ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ. ID NO 519  
 LENGTH: 1250  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (540)  
 OTHER INFORMATION: n equals a,t,g, or c  
 NAME/KEY: SITE  
 LOCATION: (1242)  
 OTHER INFORMATION: n equals a,t,g, or c  
 US-09-764-864-519

Query Match 65.3%; Score 1157.6; DB 10; Length 1250;  
 Best Local Similarity 98.8%; Pred. No. 1.3e-308; Indels 5; Gaps 4;  
 Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

QY 547 CTGTACCGCAAGCAATTTCAAGTGTCCAGAGGTTGGGAAAGCACCGTGAAGTTC 606  
 DB 18 CTGTACCGCAAGCAATTTCAAGTGTCCAGAGGTTGGGAAAGCACCGTGAAGTTC 77  
 QY 607 TTGTGTGGCCCTTATCCAAAGCGCGGGGCTGTGACCTTCCCAAGCAATCCATCTGC 666  
 DB 78 TTGTGTGGCCCTTATCCAAAGCGCGGGGCTGTGACCTTCCCAAGCAATCCATCTGC 137  
 QY 667 TTATGAGATGGGCGCTGTGAGAGCCCTCATCTTAACTTCCCGGACCCGCAATAGGT 726  
 DB 138 TTATGAGAGAGGGGCGCTGTGAGAGCCCTCATCTTAACTTCCCGGACCCGCAATAGGT 197  
 QY 727 CATCCCAATGTCTACCGGGGTGAGTACATGTGTGATCTATGCGGCAATCTGA 786  
 DB 198 CATCCCAATGTCTACCGGGGTGAGTACATGTGTGATCTATGCGGCAATCTGA 257  
 QY 787 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTCAATGCTATGATAGT 846  
 DB 258 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTCAATGCTATGATAGT 317  
 QY 847 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTCAATGCTATGATAGT 905  
 DB 318 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTCAATGCTATGATAGT 377  
 QY 906 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTTATCTTTATCTGTCTCGAAGGCTAC 965  
 DB 378 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTTATCTTTATCTGTCTCGAAGGCTAC 437  
 QY 966 GGAATGCAAGGCTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 1025  
 DB 438 GGAATGCAAGGCTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 497  
 QY 1026 TTGGAGAGGCTTCACTACGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
 DB 498 TTGGAGAGGCTTCACTACGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 557  
 QY 1085 GGAATGTCGTGTGTCATGTAATTTGATTAACCAATGATTTGGTGGATCTTTAAG 1144  
 DB 558 GGAATGTCGTGTGTCATGTAATTTGATTAACCAATGATTTGGTGGATCTTTAAG 617  
 QY 1145 TGCACCAATTTTCCATTAAGCAATGTTGACCCATGCGCTTTAAACAAGAGCTTGC 1204  
 DB 618 TGCACCAATTTTCCATTAAGCAATGTTGACCCATGCGCTTTAAACAAGAGCTTGC 677  
 QY 1205 CCATGTCATATGTGCACTACTCAAGCTTTGGGAAATGAGTGAGATGTTGAAGTGA 1264  
 DB 678 CCATGTCATATGTGCACTACTCAAGCTTTGGGAAATGAGTGAGATGTTGAAGTGA 737

QY 1265 TCAGTCTTTAACAAGTCCCTGATCCATGAATAATCTAATAGTCCCTCCCAATGA 1324  
DB 738 TCAGTCTTTAACAAGTCCCTGATCCATGAATAATCTAATAGTCCCTCCCAATGA 797  
QY 1325 GAGGATATGCGAGCGAGACGATCATCTGATGATGCTGATGATGAGGAGAAAGATGA 1384  
DB 798 GAGGATATGCGAGCGAGACGATCATCTGATGATGCTGATGATGAGGAGAAAGATGA 857  
QY 1385 CGGCTCTGAGGAGAAAGATGATCAAAATGAAGTCTAAGCTGATGAACCAATGA 1444  
DB 858 CGGCTCTGAGGAGAAAGATGATCAAAATGAAGTCTAAGCTGATGAACCAATGA 917  
QY 1445 GCAATCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1504  
DB 918 GCAATCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 977  
QY 1505 GAACTCTTAATCAAGAGACTGCTGATGAGAAATTAATCTTAATCTGATGAATAG 1564  
DB 978 GAACTCTTAATCAAGAGACTGCTGATGAGAAATTAATCTTAATCTGATGAATAG 1037  
QY 1565 AAACTTGAACATTTAGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624  
DB 1038 AAACTTGAACATTTAGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097  
QY 1625 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1684  
DB 1098 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1157  
QY 1685 TATAGTAAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1744  
DB 1158 TATAGTAAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1214  
QY 1745 AATATTTTCATAGGCAATTTTC 1767  
DB 1215 AATATTTTCATAGGCAATTTTC 1237

RESULT 3  
US-09-878-178-1815  
; Sequence 1815, Application US/09878178  
; Patent No. US20020177552A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secret, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.527  
; CURRENT APPLICATION NUMBER: US/09/878,178  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 2237  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1815  
; LENGTH: 643  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(643)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-878-178-1815

Query Match 21.0%; Score 372.6; DB 10; Length 643;  
Best Local Similarity 98.0%; Pred. No. 4.3e-92;  
Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1369 ACAGGGAACAGATGACCGCTCTGAGAGAAACAGTGCAGTCAACAATGAAGTGTACA 1428  
DB 1 ACAGGGAACAGATGACCGCTCTGAGAGAAACAGTGCAGTCAACAATGAAGTGTACA 60  
QY 1429 GCTGTAACCATGAGCAAAATCTGTGAGAGAGAGATGATTTCTCATGAGAGCAACC 1488  
DB 1429 GCTGTAACCATGAGCAAAATCTGTGAGAGAGAGATGATTTCTCATGAGAGCAACC 1488

DB 61 GCTGTAACCATGAGCAAAATCTGTGAGAGAGAGATGATTTCTCATGAGAGCAACC 120  
QY 1489 AACCTTGAAGAGAGAGAACTCTTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548  
DB 121 AACCTTGAAGAGAGAGAACTCTTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 1549 AATCTGTGTAATGAGAACTTGAACCTTTGATTAACAGAGAGAGAGAGAGAGAGAG 1608  
DB 181 AATCTGTGTAATGAGAACTTGAACCTTTGATTAACAGAGAGAGAGAGAGAGAGAG 240  
QY 1609 AGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1668  
DB 241 AGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300  
QY 1669 ATGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1728  
DB 301 ATGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 358  
QY 1729 CCACCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1774  
DB 359 CCA-CAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403

RESULT 4  
US-10-046-935-1815  
; Sequence 1815, Application US/10046935  
; Publication No. US20020156011A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secret, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.527C1  
; CURRENT APPLICATION NUMBER: US/10/046,935  
; CURRENT FILING DATE: 2002-01-15  
; NUMBER OF SEQ ID NOS: 2239  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1815  
; LENGTH: 643  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: 584  
; OTHER INFORMATION: n = A,T,C or G  
US-10-046-935-1815

Query Match 21.0%; Score 372.6; DB 13; Length 643;  
Best Local Similarity 98.0%; Pred. No. 4.3e-92;  
Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1369 ACAGGGAACAGATGACCGCTCTGAGAGAAACAGTGCAGTCAACAATGAAGTGTACA 1428  
DB 1 ACAGGGAACAGATGACCGCTCTGAGAGAAACAGTGCAGTCAACAATGAAGTGTACA 60  
QY 1429 GCTGTAACCATGAGCAAAATCTGTGAGAGAGAGATGATTTCTCATGAGAGCAACC 1488  
DB 61 GCTGTAACCATGAGCAAAATCTGTGAGAGAGAGATGATTTCTCATGAGAGCAACC 120  
QY 1489 AACCTTGAAGAGAGAGAACTCTTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548  
DB 121 AACCTTGAAGAGAGAGAACTCTTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 1549 AATCTGTGTAATGAGAACTTGAACCTTTGATTAACAGAGAGAGAGAGAGAGAGAG 1608  
DB 181 AATCTGTGTAATGAGAACTTGAACCTTTGATTAACAGAGAGAGAGAGAGAGAGAG 240  
QY 1609 AGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1668  
DB 241 AGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300



QY 1669 ATGACATATATATAGCTAATGTAATAAGCTTAAATAATTAACCTGTAATCTTTT 1728  
 DB 301 ATGACATATATATATAGCTAATGTAATA--TGGCTTAAATAATTTAAGCTGTAATCTTTT 358  
 QY 1729 CCACCAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 1774  
 DB 359 CCA-CAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 403

## RESULT 5

US-10-146-502-1815  
 ; Sequence 1815, Application US/10146502  
 ; Publication No. US20030069180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Secrist, Heather  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Stolk, John A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
 ; FILE REFERENCE: 210121.527C2  
 ; CURRENT APPLICATION NUMBER: US/10/146,502  
 ; CURRENT FILING DATE: 2002-05-14  
 ; NUMBER OF SEQ ID NOS: 2241  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1815  
 ; LENGTH: 643  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 584  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-146-502-1815

Query Match 21.0%; Score 372.6; DB 14; Length 643;  
 Best Local Similarity 98.0%; Pred. No. 4.3e-92;  
 Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1369 ACAGGAAACAGATGAAACCGCTCTGAGAGAACAGTGCAGTCAACAATGAAAGTCTCA 1428  
 DB 1 ACAGGAAACAGATGAAACCGCTCTGAGAGAACAGTGCAGTCAACAATGAAAGTCTCA 60  
 QY 1429 GCTGTAAACCAATGAAAGCAAACTTCTGAGAGAGATGTTATTCCTGATGAGAACACC 1468  
 DB 61 GCTGTAAACCAATGAAAGCAAACTTCTGAGAGAGATGTTATTCCTGATGAGAACACC 120  
 QY 1489 AACCTTGAAGAGAGCAAACTCTATCATAGAGACTGCTGTTGAGAAATTAATCTTA 1548  
 DB 121 AACCTTGAAGAGAGCAAACTCTATCATAGAGACTGCTGTTGAGAAATTAATCTTA 180  
 QY 1549 AAATCTGTAAATTAAGTAAATCTGAAACCATTAAGTAAATCAAGAACTGCAAGAGGCT 1608  
 DB 181 AAATCTGTAAATTAAGTAAATCTGAAACCATTAAGTAAATCAAGAACTGCAAGAGGCT 240  
 QY 1609 AGTTCTATTAATTAATGATTAATTTTAAATTAAGAGATGTAAGAAAGTCTCAG 1668  
 DB 241 AGTTCTATTAATTAATGATTAATTTTAAATTAAGAGATGTAAGAAAGTCTCAG 300  
 QY 1669 ATGACATATATATAGCTAATGTAATAAGCTTAAATAATTAACCTGTAATCTTTT 1728  
 DB 301 ATGACATATATATAGCTAATGTAATA--TGGCTTAAATAATTTAAGCTGTAATCTTTT 358  
 QY 1729 CCACCAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 1774  
 DB 359 CCA-CAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 403

RESULT 6  
 US-10-066-543-1696  
 ; Sequence 1696, Application US/1006543

; Publication No. US20030087818A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Pyle, Ruth A.  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Indrias, Carol Joseph  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Secrist, Heather  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Smith, Carole L.  
 ; APPLICANT: Durham, Margareta  
 ; APPLICANT: Stolk, John A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
 ; FILE REFERENCE: 210121.563  
 ; CURRENT APPLICATION NUMBER: US/10/066,543  
 ; CURRENT FILING DATE: 2002-01-31  
 ; NUMBER OF SEQ ID NOS: 3417  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1696  
 ; LENGTH: 509  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-066-543-1696

Query Match 20.7%; Score 367.4; DB 14; Length 509;  
 Best Local Similarity 97.3%; Pred. No. 9.9e-91;  
 Matches 395; Conservative 0; Mismatches 7; Indels 4; Gaps 2;

QY 1369 ACAGGAAACAGATGAAACCGCTCTGAGAGAACAGTGCAGTCAACAATGAAAGTCTCA 1428  
 DB 1 ACAGGAAACAGATGAAACCGCTCTGAGAGAACAGTGCAGTCAACAATGAAAGTCTCA 60  
 QY 1429 GCTGTAAACCAATGAAAGCAAACTTCTGAGAGAGATGTTATTCCTGATGAGAACACC 1468  
 DB 61 GCTGTAAACCAATGAAAGCAAACTTCTGAGAGAGATGTTATTCCTGATGAGAACACC 120  
 QY 1489 AACCTTGAAGAGAGCAAACTCTATCATAGAGACTGCTGTTGAGAAATTAATCTTA 1548  
 DB 121 AACCTTGAAGAGAGCAAACTCTATCATAGAGACTGCTGTTGAGAAATTAATCTTA 180  
 QY 1549 AAATCTGTAAATTAAGTAAATCTGAAACCATTAAGTAAATCAAGAACTGCAAGAGGCT 1608  
 DB 181 AAATCTGTAAATTAAGTAAATCTGAAACCATTAAGTAAATCAAGAACTGCAAGAGGCT 240  
 QY 1609 AGTTCTATTAATTAATGATTAATTTTAAATTAAGAGATGTAAGAAAGTCTCAG 1668  
 DB 241 AGTTCTATTAATTAATGATTAATTTTAAATTAAGAGATGTAAGAAAGTCTCAG 300  
 QY 1669 ATGACATATATATAGCTAATGTAATAAGCTTAAATAATTAACCTGTAATCTTTT 1728  
 DB 301 ATGACATATATATAGCTAATGTAATA--TGGCTTAAATAATTTAAGCTGTAATCTTTT 357  
 QY 1729 CCACCAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 1774  
 DB 359 CCA-CAAACTCATTAATATATTTTTCATAGGCAAGTTCCCTCAG 402

RESULT 7  
 US-09-814-353-21198  
 ; Sequence 21198, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; IDENTIFICATION, IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: NRI-0063  
 ; CURRENT APPLICATION NUMBER: US/09/814,353  
 ; CURRENT FILING DATE: 2001-03-21

```

; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21198
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-21198

```

```

Query Match      18.1%; Score 320.6; DB 12; Length 1188;
Best Local Similarity 97.8%; Pred. No. 1.5e-77;
Matches 357; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

```

```

QY 1410 CACAAATGAAAGCTACAGCTGTGTAACCAAGCAAAATTCGTGCGAGTGTGTA 1469
    |||||
DB 1 CAACAATGAAAGCTACAGCTGTGTAACCAAGCAAAATTCGTGCGAGTGTGTA 60
    |||||
QY 1470 TTCCTCATGTGACCAACCACTTTGAAAGAGCAACTCTTAATCAAGAGCTGCTG 1529
    |||||
DB 61 TTCTCATGTG-GACCAACCACTTTGAAAGAGCAACTCTTAATCAAGAGCTGCTG 119
    |||||
QY 1530 TTCGAAATTAATCTTAATTAATCTGTGTAATAGAAAACCTGACCATTAATTAACA 1589
    |||||
DB 120 TTCGAAATTAATCTTAATTAATCTGTGTAATAGAAAACCTGACCATTAATTAACA 179
    |||||
QY 1590 GAACTGCAATCAGAGGCGCTAGTCTTATTAATTAATTAATTAATTAATTAATTAATTA 1649
    |||||
DB 180 GAACTGCAATCAGAGGCGCTAGTCTTATTAATTAATTAATTAATTAATTAATTAATTA 239
    |||||
QY 1650 TGATACCTGAAAGTGTCTGATGACTTAATTAATTAATTAATTAATTAATTAATTAAT 1709
    |||||
DB 240 TGATACCTGAAAGTGTCTGATGACTTAATTAATTAATTAATTAATTAATTAATTAAT 297
    |||||
QY 1710 TTAACCTGTAACCTTTTTCACCAACATCATTAATTAATTAATTAATTAATTAATTAAT 1769
    |||||
DB 298 TTAACCTGTAACCTTTTTCACCAACATCATTAATTAATTAATTAATTAATTAATTAAT 356
    |||||
QY 1770 CTCAG 1774
    |||||
DB 357 CTCAG 361
    |||||

```

```

RESULT 8
US-09-920-300A-1176/c
; Sequence 1176, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon B.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

US-09-920-300A-1176

```

```

Query Match      17.5%; Score 310; DB 10; Length 435;
Best Local Similarity 97.4%; Pred. No. 6.2e-75;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

```

```

QY 1430 CTGGTAACCAATGAGCAAAATTCGTGCAAGTGTATTCCTCATGTCAGCAACCA 1489
    |||||
DB 435 CTGGTAACCAATGAGCAAAATTCGTGCAAGTGTATTCCTCATGTCAGCAACCA 376
    |||||
QY 1490 ACCCTTGAAGAGCAAACTCTTAATCAAGAGCTGCTGTGGAATTAATCTTAA 1549
    |||||
DB 375 ACCCTTGAAGAGCAAACTCTTAATCAAGAGCTGCTGTGGAATTAATCTTAA 316
    |||||
QY 1550 AATCTGTAAATGAAATCTGACCATTAATTAATTAATTAATTAATTAATTAATTAAT 1609
    |||||
DB 315 AATCTGTAAATGAAATCTGACCATTAATTAATTAATTAATTAATTAATTAATTAAT 256
    |||||
QY 1610 GTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1669
    |||||
DB 255 GTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 196
    |||||
QY 1670 TGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1729
    |||||
DB 195 TGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 138
    |||||
QY 1730 CACCAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1774
    |||||
DB 137 CA-CAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 94
    |||||

```

```

RESULT 9
US-10-099-926-1176/c
; Sequence 1176, Application US/10099926
; Publication No. US2003016064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon B.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-1176

```

```

Query Match      17.5%; Score 310; DB 12; Length 435;
Best Local Similarity 97.4%; Pred. No. 6.2e-75;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

```

```

QY 1430 CTGGTAACCAATGAGCAAAATTCGTGCAAGTGTATTCCTCATGTCAGCAACCA 1489
    |||||
DB 435 CTGGTAACCAATGAGCAAAATTCGTGCAAGTGTATTCCTCATGTCAGCAACCA 376
    |||||
QY 1490 ACCCTTGAAGAGCAAACTCTTAATCAAGAGCTGCTGTGGAATTAATCTTAA 1549
    |||||
DB 375 ACCCTTGAAGAGCAAACTCTTAATCAAGAGCTGCTGTGGAATTAATCTTAA 316
    |||||
QY 1550 AATCTGTAAATGAAATCTGACCATTAATTAATTAATTAATTAATTAATTAATTAAT 1609
    |||||
DB 315 AATCTGTAAATGAAATCTGACCATTAATTAATTAATTAATTAATTAATTAATTAAT 256
    |||||
QY 1610 GTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1669
    |||||
DB 255 GTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 196
    |||||

```

QY 1670 TGAATAATATGCTATAGTAAATGAGCTTAAATTAACCTGTTAACTTTTTC 1729  
DB 195 TGACTAATATATGCTATAGTAA--TGAGCTTAAATATTAACCTGTTAACTTTTTC 138  
QY 1730 CACCAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 1774  
DB 137 CA-CAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 94

## RESULT 10

US-10-033-528-1176/c  
; Sequence 1176, Application US/10033528  
; Publication No. US20020131971A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Mesgher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secret, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C1  
; CURRENT APPLICATION NUMBER: US/10/033.528  
; CURRENT FILING DATE: 2001-12-26  
; NUMBER OF SEQ ID NOS: 1896  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1176  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-033-528-1176

Query Match 17.5%; Score 310; DB 13; Length 435;  
Best Local Similarity 97.4%; Pred. No. 6.2e-75;  
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1430 CTGCTAACCATGACCAATTTCTGCGAGTGAATTTTCTCATGTMGACCAACCA 1489  
DB 435 CTGGTAAACCATGACCAATTTCTGCGAGTGAATTTTCTCATGTMGACCAACCA 376  
QY 1490 ACCCTTGAAGAGCAACCTCTCATCAAGAGCTGCTTCCAGAAATTAATCTTAA 1549  
DB 375 ACCTTTGAAGAGCAACCTCTCATCAAGAGCTGCTTCCAGAAATTAATCTTAA 316  
QY 1550 AATCTGTGTAATAGAAACCTTGAACATTAATTAATTAATTAATTAATTAATTA 1609  
DB 345 AATCTGTGTAATAGAAACCTTGAACATTAATTAATTAATTAATTAATTAATTA 256  
QY 1610 GTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1669  
DB 255 GTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 196  
QY 1670 TGACTAATATATGCTATAGTAAATGAGCTTAAATTAACCTGTTAACTTTTTC 1729  
DB 195 TGACTAATATATGCTATAGTAA--TGAGCTTAAATATTAACCTGTTAACTTTTTC 138  
QY 1730 CACCAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 1774  
DB 137 CA-CAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 94

## RESULT 11

US-09-983-965-5551  
; Sequence 5551, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Machalagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 37-21 (10297) C

; CURRENT APPLICATION NUMBER: US/09/983,965  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 5551  
; LENGTH: 387  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (347)  
; OTHER INFORMATION:  
; OTHER INFORMATION: Clone ID: 50-LIB34-010-Q1-E2-E6  
US-09-983-965-5551

Query Match 15.5%; Score 274.6; DB 10; Length 387;  
Best Local Similarity 86.8%; Pred. No. 3.4e-65;  
Matches 336; Conservative 0; Mismatches 46; Indels 5; Gaps 3;

QY 1386 CGCTTGGAGGAACAGCTGACGACCAAAATGAAGCTACAGCTGTTAACTGAGAG 1445  
DB 1 CACCTTGGAGGAACAGCTGACGACCAAAATGAAGCTACAGCTGTTAACTGAGAG 60  
QY 1446 CAATTCCTGAGCACTGAGATTAATCTCATGAGCAACCAACCTTGAAGAGAG 1505  
DB 61 CAATTCCTGAGCACTGAGATTAATCTCATGAGCAACCAACCTTGAAGAGAG 120  
QY 1506 AAATCTGTAATCAAGAGCTGCTGCGAATTAATTAATTAATTAATTAATTAATGA 1565  
DB 121 AAATCTGTAATCAAGAGCTGCTGCGAATTAATTAATTAATTAATTAATTAATGA 180  
QY 1566 AAATCTGTAATCAAGAGCTGCTGCGAATTAATTAATTAATTAATTAATTAATGA 1624  
DB 181 AAATCTGTAATCAAGAGCTGCTGCGAATTAATTAATTAATTAATTAATTAATGA 240  
QY 1625 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 1684  
DB 241 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 300  
QY 1685 TATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 1744  
DB 301 TATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 356  
QY 1745 AATATTTTCAAGAGCAAGTTCCTCT 1771  
DB 357 AATATTTTCAAGAGCAAGTTCCTCT 383

## RESULT 12

US-10-060-036-25  
; Sequence 25, Application US/10060036  
; Publication No. US20030073144A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Iodes, Michael J.  
; APPLICANT: Persing, David H.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Liang, Yudi  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060,036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Homo sapiens



|    |      |   |      |
|----|------|---|------|
| QY | 699  | TTAACTCCCGGGAAACCCGCAATAGAGTCAATCCCATCTCTCAACCCGGGTGCACTGAC | 758  |
| Db | 413  | ACAAATGAGAGACGCTACGGGAACATCACTCTGCCATCTCTCAACGGGAACAGAAAT   | 472  |
| QY | 759  | TTGTGGCAATCATTATCGCAATCTGAAAGCAAAAAATTCGCACTATTTCAAAGG      | 818  |
| Db | 473  | TATGTGATATTATATTACCTATCCAAAGAAAGAAATTTGGAGCTGTGTGAAAGG      | 532  |
| QY | 819  | GCATTAAGTACAAATGTGCTATGAAAGTAAAGAAAAACATGGCCCTTGGGTGAATCAC- | 877  |
| Db | 533  | GAATTCAGTAAACATGACCAATGGGGTTGGACCCCGCATGTACAGAGTTCTATACGG   | 592  |
| QY | 878  | --TATTCATTTTCTGTCTGTCTGTCTCTTTTATTTATTAAGGCGGCACTGTGGCT     | 935  |
| Db | 593  | GTCAGCTGTGTGTGTGTGGCCATTCGCTTCATCAACAAATGATTAATCTCGTTAGCT   | 652  |
| QY | 936  | ATTATTCCTTTTATTCGTGCGAAGGCGTAAGAAATGCAAGAGCTCAAGACGAAAGACA  | 995  |
| Db | 653  | GGCTAATATTTCATTAATACAGGCTTTCCTAATATCTGGCTCTGAATGGAAGTACA    | 712  |
| QY | 996  | GGCAATTTAAGCAGATGCTAATAAAAGCTATTGGAAGGCTTCACTACGACACTGAAC   | 1055 |
| Db | 713  | GCATAGAAA---AGAACTAAGAAAGTTTGTGGCAAGCTTCACTCACTGTAAAGC      | 769  |
| QY | 1056 | AAGGAACACAGAAATTTGGCCCTAATGSAATATGTTGGCTGTGTGCATTAATGTGATA  | 1115 |
| Db | 770  | ATGGAAGAAAAGGAATTGATGTTAATGCTAATAATTTGCAAGTGTATTAATAATTTCA  | 829  |
| QY | 1116 | AACCAATGATTTGGTACGATCTTAACTGCAACCAATATTTCCATTAAGCATGTGTG    | 1175 |
| Db | 830  | AAGTAAGGATATTATTAGAAATTCGCAACGACAGAAATTTTCCATTAATATGACATG   | 889  |
| QY | 1176 | ACCAATGCTGTAAACACAGACTTGCCCACTGTGCAGAAATGACATACCTAAAGCTT    | 1235 |
| Db | 890  | ACCAATGCTTTTGATACACGACACATGTCAATGTGTAACCTGATGTCAATCAAGCC    | 949  |
| QY | 1236 | TGGGAATTGAGTGGATGTTGAAGTGA                                  | 1264 |
| Db | 950  | TAGGATATTGGGAGACCTCGGAGATGTA                                | 978  |

```

1 NAME: Jane E. R. Potter
2 REGISTRATION NUMBER: 33,332
3 REFERENCE/DOCKET NUMBER: 1269,002
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (510) 923-2718
6 TELEFAX: (510) 655-3542
7
8 TEXT: <Unknown>
9
10 INFORMATION FOR SEQ ID NO: 1:
11
12     SEQUENCE CHARACTERISTICS:
13
14         LENGTH: 2063 base pairs
15
16         TYPE: nucleic acid
17
18         STRANDEDNESS: single
19
20         TOPOLOGY: linear
21
22     FEATURE:
23
24         SEQUENCE DESCRIPTION: SEQ ID NO: 1:
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
101
```

QY 1116 AACCAATGATTTGGTAGGCACTTTAACTGCAACCATATTTTCATAGACATGTGTG 1175  
Db 951 AAGTAAAGATATATATAGAAATCTGCATGCAAGCATATTTTCAATAGATATGCAATTG 1010  
QY 1176 ACCCATGGCTGTTAAACACAAGACTTGCCCCATGTGCAATGTGACATACCTCAAGCTT 1235  
Db 1011 ACCCATGGCTTTTGGATCACCGAACAATGTCATATGTATTAACCTTGATGTCACTAAAGCCC 1070  
QY 1236 TGGGAATTGAGTGTGANTGTTGAAGATGSA 1264  
Db 1071 TAGGATATTGGGAGAGCCTGGGGATGTA 1099

Search completed: October 6, 2003, 19:26:16  
Job time : 466 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 14:32:48 ; Search time 118 seconds  
(without alignments)  
6635,709 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774  
Sequence: 1 agctggagctccaccgcgt.....ataggcaagtttcctctcag 1774

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCOTUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                  | Description       |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1          | 254.8 | 14.4        | 1355   | 4 US-09-599-360B-64 | Sequence 64, Appl |
| 2          | 147.6 | 8.3         | 1253   | 2 US-08-786-606-6   | Sequence 6, Appl  |
| 3          | 147.6 | 8.3         | 1253   | 4 US-09-016-434-5   | Sequence 5, Appl  |
| 4          | 144.4 | 8.1         | 1251   | 4 US-09-205-258-156 | Sequence 156, App |
| 5          | 77    | 4.3         | 912    | 2 US-09-090-567-3   | Sequence 3, Appl  |
| 6          | 76.4  | 4.3         | 2605   | 2 US-08-680-395-4   | Sequence 4, Appl  |
| 7          | 75.8  | 4.3         | 2583   | 4 US-09-857-556A-9  | Sequence 9, Appl  |
| 8          | 75.6  | 4.3         | 1906   | 4 US-09-904-615-32  | Sequence 32, Appl |
| 9          | 75.6  | 4.3         | 1918   | 4 US-09-268-111-17  | Sequence 17, Appl |
| 10         | 74.6  | 4.2         | 1463   | 4 US-09-501-115-3   | Sequence 3, Appl  |
| 11         | 74.6  | 4.2         | 1774   | 4 US-09-489-847-17  | Sequence 17, Appl |
| 12         | 74.2  | 4.2         | 632    | 4 US-09-489-847-56  | Sequence 56, Appl |
| 13         | 74    | 4.2         | 762    | 4 US-09-489-847-62  | Sequence 62, Appl |
| 14         | 73.8  | 4.2         | 1378   | 1 US-08-075-533-20  | Sequence 20, Appl |
| 15         | 73.8  | 4.2         | 1378   | 2 US-08-948-176-20  | Sequence 20, Appl |
| 16         | 73.8  | 4.2         | 1378   | 5 PCT-US91-09160-20 | Sequence 20, Appl |
| 17         | 73.8  | 4.2         | 1558   | 3 US-09-043-937A-1  | Sequence 3, Appl  |
| 18         | 73.8  | 4.2         | 1558   | 3 US-09-043-937A-1  | Sequence 1, Appl  |
| 19         | 73.2  | 4.1         | 1543   | 4 US-09-227-357-125 | Sequence 125, App |
| 20         | 73    | 4.1         | 1545   | 4 US-09-227-357-125 | Sequence 25, Appl |
| 21         | 73    | 4.1         | 819    | 4 US-09-435-019-25  | Sequence 71, Appl |
| 22         | 73    | 4.1         | 921    | 4 US-09-489-847-71  | Sequence 80, Appl |
| 23         | 73    | 4.1         | 1088   | 4 US-09-489-847-80  | Sequence 64, Appl |
| 24         | 73    | 4.1         | 1517   | 4 US-09-904-615-64  | Sequence 13, Appl |
| 25         | 73    | 4.1         | 1517   | 4 US-09-345-473B-13 | Sequence 54, Appl |
| 26         | 73    | 4.1         | 1585   | 4 US-09-183-861-54  | Sequence 54, Appl |
| 27         | 73    | 4.1         | 1585   | 4 US-09-022-765-54  | Sequence 54, Appl |

|    |      |     |      |                     |                   |
|----|------|-----|------|---------------------|-------------------|
| 28 | 73   | 4.1 | 1585 | 4 US-09-551-974A-54 | Sequence 54, Appl |
| 29 | 73   | 4.1 | 2034 | 4 US-09-489-847-83  | Sequence 83, Appl |
| 30 | 73   | 4.1 | 2775 | 1 US-08-730-771-1   | Sequence 1, Appl  |
| 31 | 73   | 4.1 | 2775 | 3 US-09-060-208-1   | Sequence 1, Appl  |
| 32 | 73   | 4.1 | 3089 | 1 US-08-472-934-5   | Sequence 5, Appl  |
| 33 | 73   | 4.1 | 3089 | 2 US-08-323-460A-5  | Sequence 5, Appl  |
| 34 | 73   | 4.1 | 3089 | 2 US-08-461-146C-5  | Sequence 5, Appl  |
| 35 | 73   | 4.1 | 3089 | 3 US-08-461-146C-5  | Sequence 5, Appl  |
| 36 | 73   | 4.1 | 3089 | 4 US-08-628-829-9   | Sequence 68, Appl |
| 37 | 73   | 4.1 | 3300 | 4 US-09-482-273-68  | Sequence 1, Appl  |
| 38 | 73   | 4.1 | 3328 | 3 US-08-960-048-1   | Sequence 1, Appl  |
| 39 | 73   | 4.1 | 3328 | 4 US-09-838-586-1   | Sequence 13, Appl |
| 40 | 73   | 4.1 | 4148 | 4 US-09-435-019-13  | Sequence 15, Appl |
| 41 | 73   | 4.1 | 4148 | 4 US-09-435-019-13  | Sequence 15, Appl |
| 42 | 72.4 | 4.1 | 380  | 4 US-09-702-705-545 | Sequence 545, App |
| 43 | 72.4 | 4.1 | 380  | 4 US-09-736-457-845 | Sequence 345, App |
| 44 | 72.4 | 4.1 | 2517 | 4 US-09-857-556A-11 | Sequence 11, Appl |
| 45 | 72.2 | 4.1 | 906  | 4 US-09-489-847-72  | Sequence 72, Appl |

## ALIGNMENTS

RESULT 1  
US-09-599-360B-64  
Sequence 64, Application US/09599360B  
Patent No. 6548633  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Bouquellet, L.  
TITLE OR INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
FILE REFERENCE: GENSET.050C93  
CURRENT APPLICATION NUMBER: US/09/599,360B  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141,032  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/469,099  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patent.pm  
SEQ ID NO 64  
LENGTH: 1355  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 238..1152  
NAME/KEY: sig.peptide  
LOCATION: 238..339  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 8.50  
OTHER INFORMATION: seq SIFILSFPSNM/XA  
NAME/KEY: polyA.signal  
LOCATION: 1298..1303  
NAME/KEY: polyA.site  
LOCATION: 1324..1355  
US-09-599-360B-64  
Query Match 14.4%; Score 254.8; DB 4; Length 1355;  
Best Local Similarity 0.48; Pred. No. 8.6e-61;  
Matches 523; Conservative 0; Mismatches 312; Indels 31; Gaps 5;  
QY 449 TGGAGCTAGGAGAGAGGCGGTGACGCGAGACTGCGCGGTGAGAGCTGTGCTGG 508  
DB 397 TCGAATTGAGAGAGAGAGGTTCGCGAATCTCTCTGGAAAGGTGTCTGTGT 456  
QY 509 GTCTGTATCCCGCCGAGCGGCGCGGCGCTTAACGCTGTAAACCGACACGATTTTC 568  
DB 457 GTGTGTGCACTTCTCTGAG---AATGGAATGAGAAATGCTGTCTGTATCTTACCAATTTC 513



```

QY 569 ACGTCCCAAGGTTTGGGGAAGCACCGCTGCAAGTCTCTGGTTGGCCCTCATCCAGGC 628
| | | | |
DB 514 ACGAGGCCA-----AACAGGAGATCTATGCTGGCCCTCATCCAGAC-- 556
| | | | |
QY 629 GCGGGGGGCTGCACTTCGACAGACAAATCCATCTGGCTTATGAGAGATGGCGCTGGA 688
| | | | |
DB 557 -GTGAGGCTGTACTTTTACATTAATCAAGTGGCGAGCAAGAGGAGCAATGGG 615
| | | | |
QY 689 GCGGTCACTTTAATCTCCCGGAGACCGGCAATGAGTCAATCCCATGCTCCACCGGCT 748
| | | | |
DB 616 GTGATCATCTTACACATCTTACAGTACAGGACGATGATTTTCCCATGCTCCACAGGG 675
| | | | |
QY 749 GCAATGACATTTGTCATTCATGATCGGCAATTCGAAAGGACAAATTTCTGCACT 808
| | | | |
DB 676 ACGAAATATAGTCGGGTGATGATAGCACTGAAAGGACATGAAATTTGCACTCG 735
| | | | |
QY 809 ATTCAAGAGGCAATACAGTGAATGATGCTGATGAGTGGGAGAAAGCAATGGCCCTTGG 868
| | | | |
DB 736 ATTCAAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 795
| | | | |
QY 869 GTGATTCATCTTCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
| | | | |
DB 796 GTGACCATTTACATCATGATCTTATTTACCTTCTGCTGCTGCTGCTGCTGCTGCT 855
| | | | |
QY 929 GTGGCTATTTTATCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
| | | | |
DB 856 TTAGATGCGGTGAGAGCTTACCTTACCTGAGTGC-----CCAAATCTTCCACAGG 906
| | | | |
QY 989 AAGCAGAGCAATTAAGGCAATGCTTAAAGGCTTTTGAAGCTTCACTACGACCA 1048
| | | | |
DB 907 AGGCGAGCTCAAAATTAAGGCAATGCTTAAAGGCTTTTGAAGCTTCACTACGAGTT 966
| | | | |
QY 1049 CTGAACAAGAGACAGCAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108
| | | | |
DB 967 CTGAAGAAGGAGTGAAGATTAAGCTTAAAGGCAATGCTGCTGCTGCTGCTGCTG 1026
| | | | |
QY 1109 TTGTAAACCAATGCTTGTGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1168
| | | | |
DB 1027 ACATACAAACCCCAAGATGATGAGCACTTTTGAAGCTTGAAGCTTGAAGCTTGA 1086
| | | | |
QY 1169 TGTGTGAACCATGCTGCTTAAAGCAAGCTTGGCCCATGCTGCTGCTGCTGCTG 1228
| | | | |
DB 1087 TGCATTTGACCCCTGCTTTTGAAGCAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1146
| | | | |
QY 1229 AAGCTTGGGGAAT-TGAGGTGAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1287
| | | | |
DB 1147 AATCTTAAGAAATCTGAGAAATTTCTGAGAGTGAAGTGAAGTGAAGTGAAG 1206
| | | | |
QY 1288 ATCCATGAAATATCTAATGAGCT 1313
| | | | |
DB 1207 ATTAGATTAATGCTTATGCTT 1232
| | | | |

```

# RESULT 2

```

; Sequence 6, Application US/08786606
; Patent No. 5861495
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Coleman, Roger
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN ZINC-BINDING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786, 606
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy KJ
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0173 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-786-606-6

```

Query Match 8.3%; Score 147.6; DB 2; Length 1253;

Best Local Similarity - 55.9%; Pred. No. 5,2e-31;

Matches 301; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

```

QY 720 ATGAGTATCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
| | | | |
DB 242 AGAGGCAAGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
| | | | |
QY 780 ATCTGAAGGCAAAATTTCTGCAATCTATTCAAAGGCAATTCAGTCAATGCTCA 839
| | | | |
DB 302 AATGAGGCTAAGATATTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTT 361
| | | | |
QY 840 TGAAGTGAAGGCAAAATTCAGTCCCTTGGG--TGATCACTATTCATTTTTCGTTT 896
| | | | |
DB 362 TAGCTGTGGAATGCAATGCAACCAAGAACTTCAAGCGGTCTGATGCTTCTGCT 421
| | | | |
QY 897 CTGTGCTCTTTTATTTATTTAGCGGCACTGTGCGCTATTTATTTATTTATTTATTT 956
| | | | |
DB 422 CAATTCCTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 481
| | | | |
QY 957 GAAGCTAGCAATGCAAGACTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTA 1016
| | | | |
DB 482 AGAAGATCAGTTCACCAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 541
| | | | |
QY 1017 AAAAGCTATTTGAGGCTTCACTACGCACTGAAACAGAGAGCAAGAAATTTGCC 1076
| | | | |
DB 542 AGAAAGCAATCAGTAAATTTGACACAGCAAGTAAAGAGGATGACAGAAATTCGACC 601
| | | | |
QY 1077 CTGATGAGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136
| | | | |
DB 602 CAGATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
| | | | |
QY 1137 TCTTAAGTCAACATATTTTCCATTAAGCAATGCTGCTGCTGCTGCTGCTGCTGCT 1196
| | | | |
DB 662 TTTCTCCCTCAGAGCTATTTTCCACAAATCCGCGGTGATCCCTGCTGCTGCTGCT 721
| | | | |
QY 1197 AATCTGCCCCATGCAAGTGAATGATCACTCAAGAGCTTTGAGAAATGAGTGAAGT 1254
| | | | |
DB 722 GTACTGTCTATGCTCAACTTAATTTATTTAAGGCTTGGAAATTTGCTCAATTT 779
| | | | |

```

# RESULT 3

```

; US-09-016-434-5
; Sequence 5, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:

```

APPLICANT: Janice Au-Young  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,434  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0002 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 845-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1253 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: THPLIB01  
 CLONE: 010773  
 US-09-016-434-5

Query Match 8.3%; Score 147.6; DB 4; Length 1253;

Best Local Similarity 55.9%; Pred. No. 5.2e-31;

Matches 301; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 720 ATGAGTCATCCCATGCTCTACCCGGGTGCAAGTACATTTGCAATCATGATGCGCA 779  
 DB 242 AGGAGCCAGTACATGATCATCCAGGCACTGGAGATATTGCTGTCATGATACAG 301  
 QY 780 ATCTGAAGGCACAAAATCTGCATCTATTCAAAGAGGATCAAGTACAGATGCA 839  
 DB 302 AATTGAGGGGTAGAGATTTTGAGTTATGTGAGAGAAACATCTGTCACAAAGACAA 361  
 QY 840 TGAAGTGAAGAAAAACATGGCCCTTGGG--TGAATCACTATTCAATTTTTCGTT 896  
 DB 362 TAGCTGTGGAAGCTCAAGACGACGAAAGAACTTACGCGGTCTCTAGCTTCGTG 421  
 QY 897 CTGTGCTCTTTTATTATTACGGGCGCACTGAGGCAATTTTATCTTTATTTGCTC 956  
 DB 422 CAATATCTTTATTTTGTGATGATTTTCTTCAGCATAGGCTCATATTTCTTCATC 481  
 QY 957 GAAGGCTACGGAATGACAGCTCAAAAGCGAGAGAGCAATTTAAAGCAATGCTA 1016  
 DB 482 AGAAGATCAGGTACCAATAGACGACAGGAAACAGGCTGCTCGAGATGAGCA 541  
 QY 1017 AAAAGCTATTGGAAGGCTTCACTACGCACTGAAACAAAGAGCAAGGAATTTGCC 1076  
 DB 542 AGAAGCCATCAGTAAATTGCAACACGAGCAAGTAAAGAGGAGAACTGACC 601  
 QY 1077 CTGATGAGATAGTGTGCTGTGTCATGATGATGATTAACCAATGATTTGTCGCA 1136

DB 602 CAGACTTGATCATTTGAGAGCTGTCGATNAGAGCTATNAGAGAAATGATGTCGCA 661  
 QY 1137 TCTTAAGTGAACCAATTTTCCATAGACAGTGTGACCCATGCTGTAAACACA 1186  
 DB 662 TTCTCCCTGAGACATGTTTCCAAATCTGCTGATCTCCGCTTAAGAACAT 721  
 QY 1197 AGACTTCCCATGAGCAATGTGACATCTCAAGCTTTGGGAATGAGTGT 1254  
 DB 722 GTACCTGCTATGTGCAACTATATATTAAGAGCCCTGGAATTTGTCGCAATTT 779

#### RESULT 4

US-09-205-258-156

Sequence 156, Application US/09205258

Patent No. 6525174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1996-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,972

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,916

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 156  
LENGTH: 1251  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-205-258-156

Query Match 8.1%; Score 144.4; DB 4; Length 1251;  
Best Local Similarity 55.6%; Pred. No. 4e-30; Mismatches 236; Indels 3; Gaps 1;  
Matches 299; Conservative 0;  
QY 720 ATGAGTTCATCCCATGTCACCCGCGTGCAGTACAGTGGTTCAGTATGATGATGCGCA 779  
DB 189 AGAGCCAGTACCACTGATCCAGGCACTGAGCATATTCGTGATGATGACG 248  
QY 780 ATCTGAAGGCAACAATTCGCAATCTTAAGAAGGCACTCAAGTGAACGGTCA 839  
DB 249 AATTGAGGGTGAAGATTTTGAATTTCTGAGAAACATCTCTGACAAATGACAA 308  
QY 840 TAGAAGTAGGAAAAAAGAGCCCTTGG--GTGATCACTATTCATTTTTCGTT 896  
DB 309 TAGCTGTGAACTCGAATGCAACGAGAACTTCAGCCGCGCTCTTATGTCGTGT 368  
QY 897 CTGTGCTTTTATTTATTTAAGCGGCACTGAGGCTATTTATCTTTATTCGCTC 956  
DB 369 CATATCTTTATTTGTTTGAATGATTTATTTCTTCAAGCTCATNTCTATCTCATT 428  
QY 957 GAAGCTTACGATGCAAGACTCAAGCAGAGGCAATTAAGGCAATGCTA 1016  
DB 429 AGAAGATCAGGTACCAAAATGCAAGGCAAGGACCAAGGCTGCGAGATGCGCA 488  
QY 1017 AAAAGATTTGAAAGGCTTCAACTACGCACTGAAACAGAGAAAGAAATTTGCTC 1076  
DB 489 AGAAGCATCATGTAATTTGACACACAGACAGTAAAGAGGATGACAGAGAACTGACC 548  
QY 1077 CTGATGAGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136  
DB 549 CAGACTTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608  
QY 1137 TCTTACGTCGCAACCATTTTTCATTAAGACATGCTTGAACCATGCTGTTAAACACA 1196  
DB 609 TTCTCCCTGCAAGCATGTTTTCACAAATCTGCGGATCCCTGCTTGAAGAACATT 668  
QY 1197 AGACTTGCCCATGTCGAATGTGACATACCAAGCTTGGGAATGAGGTGAGT 1254

DB 669 GTACCTGTCTATGTCGAACCTTAATATTTGAAGGCGCTGGAATGTGCGCAATTT 726

RESULT 5  
US-09-090-567-3  
Sequence 3, Application US/09090567  
Patent No. 5989549  
GENERAL INFORMATION:  
APPLICANT: Sullivan, Robert  
APPLICANT: Brub, Bruno  
APPLICANT: Igar, Christine  
APPLICANT: Gaudreault, Christian  
TITLE OF INVENTION: Acrosomal Sperm Protein And  
TITLE OF INVENTION: Uses Thereof  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swabey Ogilvy Renault  
STREET: 1600 - 1981 McGill College  
CITY: Montreal  
STATE: QC  
COUNTRY: Canada  
ZIP: H3A 2Y3  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/090,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Kevin P  
REGISTRATION NUMBER: 26,674  
REFERENCE/DOCKET NUMBER: 13045-2"US" FC/CC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 514-845-7126  
TELEFAX: 514-288-8389  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 912 base pairs  
TYPE: nucleic acid  
STRANDS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
US-09-090-567-3  
Query Match 4.3%; Score 77; DB 2; Length 912;  
Best Local Similarity 94.1%; Pred. No. 1.8e-11;  
Matches 80; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 AGCTGAGCTCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60  
DB 7 AGCTGAGCTCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 66  
QY 61 AATTGCGCAGAGCCGAGAGCTGC 85  
DB 67 AATTGCGCAGAGCCGAGAGCTGC 91  
RESULT 6  
US-08-680-395-4  
Sequence 4, Application US/08680395  
Patent No. 5892010  
GENERAL INFORMATION:  
APPLICANT: Gray, Joe W.  
APPLICANT: Collins, Colin

```

NUMBER OF SEQ ID NOS: 34
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 2583
TYPE: DNA
ORGANISM: Glycine max
US-09-556A-9

Query Match
Best Local Similarity 87.4%; Score 75.6; DB 4; Length 2583;
Matches 83; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGCTGAAGCTTCACCCGCGTGGCGCGCTCTAGAACTAGTGGATGCCCGGCGTGGAGG 60
DB 46 AGCTGAAGCTTCACCCGCGTGGCGCGCTCTAGAACTAGTGGATGCCCGGCGTGGAGG 105

QY 61 AATTCGGCAGACCGAGAGATGCTATGCGGCA 95
DB 106 AATTCGGCAGACCGAGAGATGCTATGCTTCTCTCA 140

RESULT 8
US-09-904-615-32
Sequence 32, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 32
LENGTH: 1906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (617)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (940)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1461)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1801)
OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-32

Query Match
Best Local Similarity 95.1%; Score 75.6; DB 4; Length 1906;
Matches 78; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CTGGAAGCTTCACCGCGTGGCGCGCTCTAGAACTAGTGGATGCCCGGCGTGGAGG 62
DB 2 CTGGAAGCTTCACCGCGTGGCGCGCTCTAGAACTAGTGGATGCCCGGCGTGGAGG 61

QY 63 TTGGCAGACCGAGAGATG 84
DB 62 TTGGCAGACCGAGAGATG 83

RESULT 9
US-09-268-311-17

```

```
/ Sequence 17, Application US/09268311
; Patent No. 6482923
; GENERAL INFORMATION:
; APPLICANT: Shi, Yangu
; APPLICANT: Ruben, Steve M.
; TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
; FILE REFERENCE: PF398PI
; CURRENT APPLICATION NUMBER: US/09/268,311
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 60/059,133
; EARLIER FILING DATE: 1997-09-17
; EARLIER APPLICATION NUMBER: 09/154,219
; EARLIER FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1918
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111)..(1409)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (101)
; OTHER INFORMATION: n equals a, t, g or c
; US-09-268-311-17

Query Match      4.3%; Score 75.6; DB 4; Length 1918;
Best Local Similarity 89.0%; Pred. No. 6.8e-11;
Matches 81; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 60
DB 36 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 95

QY 61 AATTGGGACGAGCGGAGAGCTGATCTGC 91
DB 96 AATTGGGACGAGCGGAGTGTGCTGCTGC 126

RESULT 10
US-09-501-115-3
; Sequence 3, Application US/09501115
; Patent No. 6552249
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fader, Gary M.
; APPLICANT: Rafaleki, Antoni
; TITLE OF INVENTION: Plant Glnamyl-Alcohol Dehydrogenase Homologs
; FILE REFERENCE: BM1328 US NA
; CURRENT APPLICATION NUMBER: US/09/501,115
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 60/119,585
; EARLIER FILING DATE: 1999-February-10
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-501-115-3

Query Match      4.2%; Score 74.6; DB 4; Length 1463;
Best Local Similarity 95.1%; Pred. No. 1.1e-10;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 60
DB 141 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 200
```

```
DB 41 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 100
QY 61 AATTGGGACGAGCGGAGAG 81
DB 101 AATTGGGACGAGCGAGAG 121

RESULT 11
US-09-489-847-17
; Sequence 17, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1774
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-489-847-17

Query Match      4.2%; Score 74.6; DB 4; Length 1774;
Best Local Similarity 95.1%; Pred. No. 1.2e-10;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 60
DB 141 AGCTGAGCTCCACGCGGCTGTCGCTGAGTCTGATCCCGGCGCTGACAG 200

QY 61 AATTGGGACGAGCGGAGAG 81
DB 201 AATTGGGACGAGCGGCTGCTGC 221

RESULT 12
US-09-489-847-56
; Sequence 56, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
```

```
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 56
/ LENGTH: 632
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURES:
/   NAME/KEY: SITE
/   LOCATION: (29)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
/     NAME/KEY: SITE
/     LOCATION: (46)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
/     NAME/KEY: SITE
/     LOCATION: (94)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
/     NAME/KEY: SITE
/     LOCATION: (162)
/   OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-56

Query Match
Best Local Similarity 96.2%; Score 74.2; DB 4; Length 632;
Best Local Similarity 96.2%; Pred. No. 9e-11;
Matches 76; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGCTGAGCTCCACCGCGGCGGCTCTGAGACTAGTGATCCCGCGGCTGCAGG 60
DB 167 AGCTGAGCTCCACCGCGGCGGCTCTGAGACTAGTGATCCCGCGGCTGCAGG 225
OY 61 AATTGGCAGCAGCGGAGG 79
DB 227 AATTGGCAGCAGCGATGATG 245

RESULT 13
US-09-489-847-62
/ Sequence 62, Application US/09489847
/ Patent No. 6476195
/ GENERAL INFORMATION:
/   APPLICANT: Rosen et al
/   TITLE OF INVENTION: 98 Human Secreted Proteins
/   FILE REFERENCE: P2031P1
/   CURRENT APPLICATION NUMBER: US/09/489,847
/   EARLIER FILING DATE: 2000-01-24
/   EARLIER APPLICATION NUMBER: PCT/US99/11130
/   EARLIER FILING DATE: 1999-07-29
/   EARLIER APPLICATION NUMBER: 60/094,657
/   EARLIER FILING DATE: 1998-07-30
/   EARLIER APPLICATION NUMBER: 60/095,486
/   EARLIER FILING DATE: 1998-08-05
/   EARLIER APPLICATION NUMBER: 60/096,319
/   EARLIER FILING DATE: 1998-08-12
/   EARLIER APPLICATION NUMBER: 60/095,454
/   EARLIER FILING DATE: 1998-08-06
/   EARLIER APPLICATION NUMBER: 60/095,455
/   EARLIER FILING DATE: 1998-08-06
/   NUMBER OF SEQ ID NOS: 376
/   SOFTWARE: PatentIn Ver. 2.0
/   SEQ ID NO 62
/   LENGTH: 762
/   TYPE: DNA
/   ORGANISM: Homo sapiens
/   FEATURES:
/     NAME/KEY: SITE
/     LOCATION: (10)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
/     NAME/KEY: SITE
/     LOCATION: (12)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
```

```
/ NAME/KEY: SITE
/ LOCATION: (42)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/   NAME/KEY: SITE
/   LOCATION: (219)
/   OTHER INFORMATION: n equals a,t,g, or c
/   FEATURE:
/     NAME/KEY: SITE
/     LOCATION: (747)
/   OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-62

Query Match
Best Local Similarity 93.9%; Score 74; DB 4; Length 762;
Best Local Similarity 93.9%; Pred. No. 1.1e-10;
Matches 77; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 AGCTGAGCTCCACCGCGGCGGCTCTGAGACTAGTGATCCCGCGGCTGCAGG 60
DB 293 AGCTGAGCTCCACCGCGGCGGCTCTGAGACTAGTGATCCCGCGGCTGCAGG 352
OY 61 AATTGGCAGCAGCGGAGG 82
DB 353 AATTGGCAGCAGCAATATC 374

RESULT 14
US-08-075-533-20
/ Sequence 20, Application US/08075533
/ Patent No. 5530186
/ GENERAL INFORMATION:
/   APPLICANT: Hitz, William D.
/   APPLICANT: Yadav, Narendra S.
/   TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
/   NUMBER OF SEQUENCES: 22
/   CORRESPONDENCE ADDRESS:
/     ADDRESSEE: E. I. du Pont de Nemours and Company
/     STREET: 1007 Market Street
/     CITY: Wilmington
/     STATE: Delaware
/     COUNTRY: U.S.A
/     ZIP: 19898
/   COMPUTER READABLE FORM:
/     MEDIUM TYPE: Floppy disk
/     COMPUTER: IBM PC compatible
/     OPERATING SYSTEM: PC-DOS/MS-DOS
/     SOFTWARE: PatentIn Release #1.0, Version #1.25
/   CURRENT APPLICATION DATA:
/     APPLICATION NUMBER: US/08/075,533
/     FILING DATE:
/     CLASSIFICATION: 800
/     PRIOR APPLICATION DATA:
/       APPLICATION NUMBER: US 07/631,264
/       FILING DATE: 20-DEC-1990
/       ATTORNEY/AGENT INFORMATION:
/         NAME: Morrissey, Bruce W.
/         REGISTRATION NUMBER: 30,663
/         REFERENCE/DOCKET NUMBER: CR-8926-A
/       TELECOMMUNICATION INFORMATION:
/         TELEPHONE: (302) 992-4927
/         TELEFAX: (302) 892-7949
/         TELEX: 835420
/       INFORMATION FOR SRO ID NO: 20:
/         SEQUENCE CHARACTERISTICS:
/           LENGTH: 1378 base pairs
/           TYPE: nucleic acid
/           STRANDEDNESS: single
/           TOPOLOGY: linear
/         MOLECULE TYPE: DNA (genomic)
/         ORIGINAL SOURCE:
/           ORGANISM: Brassica napus
US-08-075-533-20
```

## Query Match

4.2%; Score 73.8; DB 1; Length 1378;

Best Local Similarity 97.4%; Pred. No. 1.8e-10;

Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Search completed: October 6, 2003, 14:43:49  
Job time : 122 secs

QY

1 AGCTGAGCTCCACCGCGGTGCGCGCTCTAGAACTAGTGAATCCCCCGGCTGCAAG 60  
|||||

DB

52 AGCTGAGCTCCACCGCGGTGCGCGCTCTAGAACTAGTGAATCCCCCGGCTGCAAG 111  
|||||

QY

61 AATGCGCAGAGCCGA 77  
|||||

DB

112 AATGCGCAGAGAGGA 128  
|||||

## RESULT 15

US-08-948-176-20

; Sequence 20, Application US/08948176

; Patent No. 5945585

; GENERAL INFORMATION:

; APPLICANT: HITZ, WILLIAM D.

; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES

; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT

; TITLE OF INVENTION: OIL COMPOSITION

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/948,176

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/631,264

; FILING DATE: DECEMBER 20, 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: CRISTENBURY, LYNN M.

; REGISTRATION NUMBER: 30,971

; REFERENCE/DOCKET NUMBER: CR-8926-C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-5481

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1378 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Brassica napus

; US-08-948-176-20

## Query Match

4.2%; Score 73.8; DB 2; Length 1378;

Best Local Similarity 97.4%; Pred. No. 1.8e-10;

Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 AGCTGAGCTCCACCGCGGTGCGCGCTCTAGAACTAGTGAATCCCCCGGCTGCAAG 60  
|||||

DB

52 AGCTGAGCTCCACCGCGGTGCGCGCTCTAGAACTAGTGAATCCCCCGGCTGCAAG 111  
|||||

QY

61 AATGCGCAGAGCCGA 77  
|||||

DB

112 AATGCGCAGAGAGGA 128  
|||||